Vaasan yliopisto
UNIVERSITY OF VAASA

Sushil Bhusal

Data-driven Predictive Maintenance for Li-Ion Batteries in Autonomous Systems using Machine Learning

School of Technology and innovations
Master of Computing Science
Sustainable and Autonomous Systems

Vaasa 2026

UNIVERSITY OF VAASA
School of Technology and Innovations
Author: Sushil Bhusal
Title of the Thesis: Data-driven Predictive Maintenance for Li-Ion Batteries in Autonomous Systems using Machine Learning
Degree: MSc in Computing Science (Technology)
Programme: Sustainable and Autonomous Systems
Supervisor: Mohammed Elmusrati
Thesis Evaluator: Dr. Elham Ahmadi
Year: 2026 **Pages:** 87

ABSTRACT:

Autonomous Mobile Robots (AMRs) in warehouse logistics, manufacturing, and industrial automation applications are powered by Li-Ion batteries. The smooth operation of this system relies on battery performance. Traditional maintenance approaches, such as run-to-failure and fixed-schedule preventive maintenance, are ineffective at addressing the progressive, non-linear capacity loss observed in Li-Ion batteries. A data-driven predictive maintenance approach that accurately predicts batteries' State of Health and Remaining Useful Life in real time provides a more efficient and environmentally friendly solution. In this thesis, such a framework is developed and tested using publicly available data from the NASA Prognostics Centre of Excellence (PCoE) battery dataset.

Two machine learning models were trained. A Random Forest regression model was used as interpretable baseline and a two-layer Long Short-Term Memory (LSTM) neural network model to account for the time-dependent nature of battery degradation data. Both models predicted SoH instead of RUL directly. RUL was then calculated from the predicted SoH by determining when the predicted SoH is projected to reach 80% end-of-life threshold. The performance of the two models on the test battery B0018 shows that the LSTM model significantly outperformed the Random forest on all four metrics (RMSE, MAE, MAPE, and R^2). The result translates to 78% improvements in RMSE, and demonstrates that the sequential modelling of discharge cycle history can delivery an accuracy gain over feature-based ensemble models.

The feature importance analysis highlight the discharge capacity as the most significant predictor, with a relative importance of 75.4% in line with the electromechanical knowledge that capacity fade is the main microscopic effect of lithium loss through solid electrolyte inter phase growth. This thesis result shows that a simple LSTM model architecture can be trained to achieve near perfect RUL prediction accuracy on a previously unseen battery with a very simple computer setup and Python. These findings set a practical, repeatable and academically rigorous approach to data-driven predictive battery maintenance for autonomous system, which will benefit the sustainability objectives of avoiding unplanned downtime, increasing lifespan, and reducing e-waste in industrial AMR fleets.

KEYWORDS: Predictive maintenance, remaining useful life, state of health, lithium-ion batteries, machine learning, Random Forest, Long Short-Term Memory, autonomous mobile robots

Contents

Pictures

Figures

Tables

Abbreviations

AMR	Autonomous Mobile Robot
BMS	Battery Management System
EIS	Electrochemical impedance spectroscopy
EOL	End-of-Life
LFP	Lithium-Ion Phosphate
LIB	Lithium-Ion Battery
LSTM	Long short-Term Memory
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
MDI	Mean Decrease in Impurity
ML	Machine Learning
MSE	Mean Squared Error
NASA	National Aeronautics and Space administration
NMC	Nickel Manganese Cobalt Oxide
PCoE	Prognostics Centre of Excellence
RF	Random Forest
RMSE	Root Mean Square Error
RNN	Recurrent Neural Network
RUL	Remaining Useful Life
SEI	Solid Electrolyte Inter phase
SoC	State of Charge
SoH	State of Health

Chapter 1: Introduction

1.1 Background and motivation

Autonomous Mobile Robots (AMR) use has increased significantly in the industrial and commercial sectors in the last decade. Fleets of battery-powered AMRs are being increasingly utilized to do different task such as to handle material, conduct inspection and deliveries in warehouses, manufacturing floors, hospitals, and logistics centres continuously without any break. These systems have great benefits over manual labouring speed, uniformity and cost of operation. But their constant and dependable functioning is fully depended on Lithium-Ion (Li-Ion) battery.

Li-Ion batteries are the most common energy storage technology used in AMRs platforms because they store large amounts of energy, the relative rate of self-discharge is lower, and cycle life is longer than other older battery technologies. Even though they have these benefits they also have demerits. Over multiple charge and discharge cycles, Li-Ion batteries experience slow, irreversible capacity degradation. The combination of electrochemical ageing processes such as lithium plating, solid electrolyte inter phase (SEI) layer and loss of active material at electrode surfaces cause this degradation. In the long run, these processes lower the usable capacity of the battery, reduce its runtime on a single charge and also cause end of life eventually when the battery is not able to provide enough power to sustain the operational requirement.

In the case of AMRs operators, battery degradation causes two impacts one is direct operational risk and another one is financial risk. A sudden failure of the battery during the working process will stop the movement of a robot on the warehouse floor and can interfere with the process of automated work, as well as cause expensive manual interventions. On the other hand, if we are unable to know the exact condition of the battery and change it too soon before they have truly reached their end of useful life wastes resource, and creates unwanted electronic waste.

Conventional battery care methods are of two general types: run-to-failure, where batteries may be operated until they fail to work and then replaced in reactive manner. And another one is predictive maintenance, where batteries are replaced on regular schedule whether they have actually degraded or not. These two methods are not so efficient. run to failure puts the whole operation at risk of unexpected downtime, whereas scheduled replacement does not care weather the batteries are working condition or not if the time has came they are replaced , although batteries still have a considerable portion of useful life remaining. There is a need of a more intelligent, data-driven solution that can track the health of specific batteries in real time, estimates the amount of useful life left in them, and does not initiate any interventions merely in response to the maintenance needs. This is referred to as AI-based predictive maintenance.

1.2 Machine learning in Battery prognostics

Remaining Useful Life (RUL) of Li-Ion battery is very difficult to predict. The degradation of batteries is a complex interaction of factors such as operating temperature, charge and discharge rate, depth of discharge and cell chemistry. Physical degradation models are precise in the controlled laboratory environment, but needs a lot of detailed information about the electrochemical parameters that are hard to measure in practice and are computationally costly to model in real time.

In contrast with data-driven approaches, data-driven strategies acquire degradation information based on past sensor measurements without necessarily having explicit knowledge of the underlying physics. Time-series data on battery management system such as voltage, current, temperature, and discharge capacity measurements can be fed to machine learning models to detect some minor patterns that are hard to see while attempting to predict capacity fed but can indicate when a battery has reached end-of-life. The models can then be implemented in real operational AMRs fleets to enable continuous and automatic health monitoring and maintenance scheduling.

Mainly there are two machine learning techniques that can be used to predict the health of batteries which are especially relevant. Random Forest regression is an established ensemble approach, which yields interpretable scores of feature importance, allowing engineers to determine the most predictive measurable quantities in battery health. Long Short-Term Memory (LSTM) neural networks, a particular kind of recurrent neural network, are optimally adapted to sequential data with long-term temporal dependencies which is an attribute that allows them to capture the history-dependent and gradual nature of battery degradation. A comparison of these two methods offer more than a useful evaluation of the relative accuracy of the two methods and a deeper view of the degradation signals that are identified by this methods.

1.3 Research Problem

Although there is increasing evidence on data-driven battery prognostics, a number of issues have been experienced in their implementation in practice. Most published research train and test their models on the same batch of experiments, or the same batch of batteries, thus restricting the evidence they can demonstrate about generalizing their results to unseen batches. Second, publicly available battery datasets, in which cell across multiple experimental conditions are co-located in storage, are heterogeneous, and thus necessitate close data quality screening that is typically not described in detail. Third, another problem that is widely encountered but ignored is the mismatched of scales that occurs when training and test batteries differ in total lifetimes by a significant fraction, leading direct RUL prediction model to make systematic over or under-prediction of the remaining life.

This thesis considers these issue by systematically and end-to-end executing a predictive maintenance structure with actual NASA battery information, specifically, cross-battery generalization, screening of prediction data, and the methodological selection of prediction target. The main research question that this thesis deal with is :

Are machine learning models, namely, Random Forest regression and LSTM neural networks, able to predict the State of Health and Remaining Useful Life of Li-Ion batteries using data on discharge cycle sensor, and generalize to a battery that was not present when the model was trained?

1.4 Research Objectives

This thesis aims at achieving three particular goals which are directly based on the above research problem:

- Analyze publicly available battery ageing data of NASA Prognostics Centre of Excellence (PCoE) dataset and determine the main characteristics that define the process of battery ageing based on the raw discharge cycle data.
- Train and develop two machine learning predictors; a Random Forest regressions and an LSTM neural network, to predict Battery State of Health (SoH) and Remaining Useful Life (RUL), and test their performance on a held-out test battery that was not considered during training.
- Compare the output of each of the two models using four evaluation measures, namely, RMSE, MAE, MAPE and R^2 and compare the feature importance output of the two models to determine the physical drivers of battery degradation that the model capture.,

1.5 Scopes and Limitations

The paper is limited to 18650 cylindrical Li-Ion battery cells that are controlled in the laboratory, as NASA PCoE provides. No battery test or hardware testing were done. All analysis is done based on a secondary data. Only discharge cycle data is taken into consideration in the study, and Electro mechanical Impedance Spectroscopy (EIS) measurements are not considered, but these are also present in dataset, and feature extraction is more complex.

The predictive maintenance system that will be created as a part of this thesis is meant to be a demonstration of the proof-of-concept of the data-driven approach. It is not tested in a real world AMR deployment scenario. How well the results can be generalized to batteries of other chemistries or form factors or operating conditions remains an open question.

The analysis is also limited to a train/test split at the battery level with three training cells and one test cell. This is in line with the typical benchmark configuration found in battery prognostics literature. Although this constrains the statistical strength of the analysis, it is a premeditated decision, which values methodological rigor.

1.6 Researach Methods and Tools

This paper is a quantitative, data-driven research. The whole analysis workflow was run in Python 3.10 using Jupyter Notebook as the interactive development and running on a typical windows laptop with the help of Anaconda. Only local CPU was used to train all models with no cloud computing resources or GPUs needed. The most important tool used is summarized below:

- SciPy: to read the NASA.mat binary data files.
- Pandas and NumPy: to manipulate data, feature engineer, and work with arryas.

- Scikit-learn: to use the Random Forest model and Min-Max normalization, as well as evaluation metrics.
- TensorFlow/ Keras: to create and train the LSTM neural network.
- Maatplotlib: to produce all the figures and visualization.

Use of open-source guarantees that all the results are completely reproducible and that framework can be replicated or even extended by other researcher even without access to proprietary software. The entire Jupyter Notebook with all of the code, outputs, and figures is a supplement deliverable of the current thesis alongside the written report.

1.7 Significance of the study

This thesis contributes to the area of battery prognostics and predictive maintenance of autonomous system in a number of ways.

- Practically, the LSTM model created in this paper return a result of R^2 = 0.993 and mean absolute error of only 3.1 cycles on a totally unknown test battery- a result that would be directly translated to a predictable maintenance schedule by AMR fleet operators. With this level of accuracy an early notification of battery end-of-life would enable maintenance personnel to plan battery replacements in advance to prevent downtime surprises as well as unnecessary replacement.
- Methodologically, this thesis records an empirical problem, which has often been recognized but rarely directly addressed in literature : the mismatch between cross-battery RUL scales that makes direct RUL prediction models fail when training and testing batteries differ in total lifetime. The conversion strategy of SoH to RUL derived here offers a straightforward and efficient solution that enhances generalization without additional data.

- Predictive maintenance with accuracy is a direct support of the objective of increasing battery service life, in terms of sustainability. By swapping batteries when they actually are nearing end-of-life, as opposed to swapping batteries based on specific timeframe, the framework created as part of this thesis can decrease the amount of LI-Ion battery waste disposed too soon, which is part of the overall sustainability goals of organization implementing autonomous systems.

1.8 Thesis Structure

The remaining part of the thesis will have following structure:

- **Chapter 2 :** Fundamentals of Li-Ion batteries
- **Chapter 3:** : AI-based Battery Predictive Maintenance Literature Review.
- **Chapter 4 :** Proposed methods and results
- **Chapter 5 :** Conclusions

Chapter 2: Fundamentals of Li-Ion Batteries

2.1 Introduction.

The excessive use of Autonomous Mobile Robots (AMR) in industrial and logistic sectors has established a requirement of high-density energy storage system that can be relied upon. The most predominent battery technology used is Lithium-Ion batteries because of its best combination of specific energy, cycle life, and charge efficiency in comparison to the old technology of lead-acid and nickel-metal hydride system (Deng, 2015. LIBs have inherent drawback such as irreversible electrochemical ageing, which is responsible for gradual decreasing of he capacity and power deliver ability of the LIBs with time. In order to make predictive maintenance work, the battery should not be regarded as a black box that merely provides a capacity value. The reflection should be done on the true knowledge of the inner processes that regulate battery behaviour .This is the key factor for making sense of sensor data, choosing meaningful features , and framing the prediction problem appropriately.

The main focus of this chapter is to lay the foundation of that electromechanical basis. Section 2.2 explains the basic principle of operation of Li-Ion cells. Section 2.3 explains the main internal elements and the way their material characteristics affects the performance and durability. In section 2.4, the operational metrics to describe battery health in real systems are defined. Section 2.5 gives an in-depth discussion of the main degradation processes as the physical and chemical mechanism that lead to capacity decay and resistance increase. In section 2.6, the State of Health (SoH) and Remaining Useful life (RUL) are formally defined and how they relate to the internal ageing processes mentioned above.

2.2 Principle of operation of Lithium-Ion Batteries.

The working mechanism of a Li-Ion cell is an electromechanically reversible process, called the rocking -chair mechanism (Deng, 2015). Li-Ion battery, in contrast to primary batteries where electrodes undergo a uni-directional chemical reaction, host electrode materials such as layred metal oxides (at the cathode) and graphite (at the anode) that enable lithium ions to be inserted and extracted through their crystal lattice structure without permanent structural damage under ideal conditions

An external voltage is applied during charging a non-spontaneous electrochemical re-action. The lithium ions move out of the cathode material, through the liquid electro-lyte, micro porous polymer separator and this is intercalate into the layered graphite structure at the anode. At the same time, an electron emitted at the cathode flows to the anode in the external circuit, keeping the charge balance. Half-reaction at cathode during charging process can be modelled as:

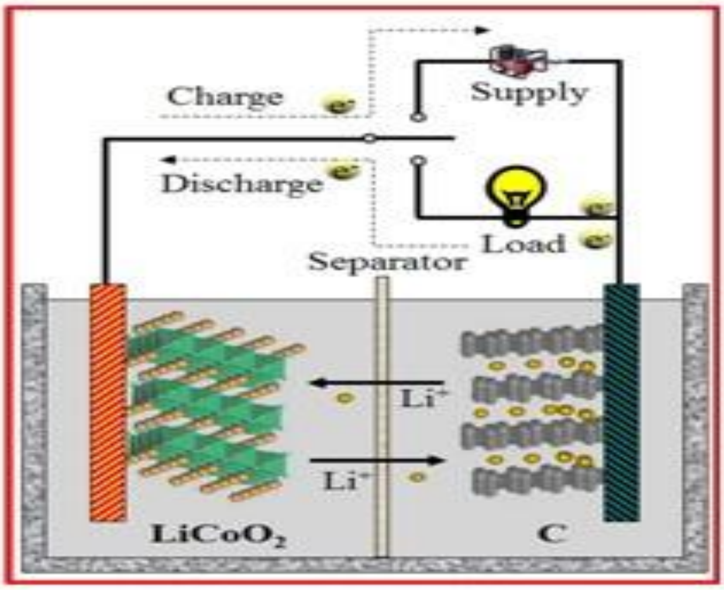
$LiMO_2 \rightarrow Li_{1-x}MO_2 + xLi^+ + xe^-$

M=the transition metal (cobalt, manganese, nickel, or a mixture of these)
x= the fraction of lithium extracted.

The intercalation reaction at anode is:

$xLi^+ + xe^- + C_6 \rightarrow Li_xC_6$

During discharge as the AMR is actively taking power to run its motors, sensors and onboard computing systems, both reactions are inverted. The Lithium ions move in reverse to the anode and cathode. The electrons move through the external circuit providing electric power to the load. The voltage at the terminals is a direct result of the difference in the electrochemical potential between the two electrodes and decreases with the loss of state of charge and accumulation of internal resistance with ageing.



Picture 2.1 Rocking-chair mechanism of lithium-ion cell (Deng, 2015)

Due to the reversibility feature Li-Ion batteries are rechargeable but also have its weakness. Each cycle exerts mechanical stress on the host lattice structures with each insertion and extraction event, and each cycle reveals the surfaces of the electrolyte and electrodes to reactive environments which causes parasitic side reaction. when thousands of cycle experience these effects there occurs visible effects of capacity fade and power fade which is characterize as battery ageing.

2.3 Internal Parts of Lithium-Ion battery and their functions.

The four main components of li-Ion cell are as follows:

1. The cathode
2. The anode
3. The electrolyte
4. The Separator

These components determine the performance, safety and durability of the cell. It is important to understand the role and constrains of all of them in order to interpret the operating conditions that promote degradation and reason for that.

2.3.1 Cathode

The positive electrode is the cathode, and the main factor of energy density and operating voltage of the cell. In commercial cells, the cathode is a lithium transition-metal oxide powder, fixed to current collector of aluminium.

The chemical composition that is chosen in warehouse AMRs and stationary storage applications is Lithium Iron Phosphate ($LiFePO_4$). It has a nominal voltage of about 3.2 V and it provides superior thermal stability. The strong phosphate bonds do not allow oxygen to escape even during thermal stress, eliminating the possibility of thermal runway precisely. Lithium Iron Phosphate has a highly flat profile of discharge voltage which results in making estimation of SoC easy. However there may be complication in SoH assessment because a small loss in capacity will lead to small changes in the shape of the discharge curve. it has a cycle life of about 2,000 and 80% SoH which is the reason it is suitable for AMR fleets which are used continuously.

Compared to Lithium Iron Phosphate, Nickle Manganese Cobalt (NMC) oxides have higher energy density but they have lower thermal stability. The NMC cells are mainly used in electric cars and high-performance AMRs where the run time on a charge has a significant limitation. NMC has a steeper discharge curve, thus is more susceptible to the voltage-based SoH estimation procedures

2.3.2 Anode

The commercial standard anode material used in NASA PCoE battery cells in this study is graphite. It has a stable hexagonal structure permitting lithium intercalation with a maximum of one lithium ion per six carbon atoms, which gives a theoretical specific capacity of 372 mAh/g.

The first cycle of charging has a critical but irreversible reaction at the surface of the graphite. The thermodynamically unstable electrolyte at low potentials of a charged graphite anode is decomposed and deposited as a thin, ion-permeable layed on the anode surface. The solid Electrolyte Inter phase (SEI) is a fundamental layer in the operation of long term degradation of Li-Ion cells.

2.3.3 Electrolyte

Electrolyte is the ionic conductor between the two electrodes. In commercial Li-Ion cells usually a lithium salt- most often lithium hexafluorophosphate ($LiPF_6$) dissolved in a solution of organic carbonate solvents like ethylene carbonate (EC) and Di-methyl carbonate (DMC). The electrolytes should be highly conductive of ion, have low electronic conductivity, have a broad electrochemical stability range and should be viscous enough that it does not become viscous over its operating temperature range.

The organic solvents of commercial electrolytes are flammable and ractive which has been the main issue for the safety in Li-Ion cells. In high voltages or high temperatures, the product of decomposition of electrolyte results in gas production, and can initiate the exothermic chai reaction of thermal runaway. The active lithium is degraded gradually by the electrolyte decomposition at both electrodes surfaces, and the resistance within cell increases with life of cell.

2.3.4 Separator

The separator is a polymer membrane, which is micro-porous, and is usually made up of polyethylene (PE) or polypropylene (PP) whose place is between the cathode and the anode. Its main application is to avoid direct electronic contact between the electrodes but it gives permission to ion movement in its pores. In modern separators, a thermal shutdown system is introduced where at high temperature the polymer melts and seals the holes, stopping the ionic current and preventing the further heat production in case of abuse. The integrity of the separator is of utmost importance to safety. If the cells are repeatedly charged at low temperature the lithium will pierce the separator which results in internal short circuit

2.4 Battery Health Monitoring Operational Metrics.

- **Discharge Capacity (Ah) :** It is the amount of charge which is discharge d in one cycle of discharge. It is the nearest measurable change of battery health as it represents joint outcome of all internal ageing processes. A new 18650 Li-Ion cell of the kind utilized in NASA PCoE dataset, has a nominal capacity of about 2.0 Ah . However, this value decreases monotonically with the cycle life as active lithium and block ion pathways are degraded.

- **Terminal Voltage:** It is the potential difference that is observed at the end of the battery when it is in running. The current and internal resistance (IR) combine to cause the difference between the terminal voltage and the open-circuit voltage (OCV). The more the battery gets older the more internal resistance increases. If the voltage drop is higher and the lower terminal voltage reaches the lower cut-off threshold the discharge cut off in the discharge will be earlier, thus limiting the usable capacity per discharge.

- **Temperature:** It is indicative of the heat lost due to internal resistance and electrochemical reactions. The increase in operating temperature results in increase of the rate of nearly all degradation processes and very low temperature decreases the iconic conductivity and increase the probability of lithium plating. Temperature monitoring is an important input in health models in AMR deployments since temperature in various areas of the warehouse and seasonal factors vary.
- **Internal Resistance:** It is a measure of the impedance to current in the cell, and includes electronic resistance of electrodes and current collectors, iconic resistance of electrolyte, and interfacial resistance of SEI and cathode-electrolyte interface. It rises with age, and is directly proportional to the power loss seen in old batteries. Although it is not explicitly recorded in the NASA PCoE discharge cycle data that is used in this thesis, it is partially reflected in features of voltage drops.
- **Charge and Discharge rate (C-rate) :** It is the current as a fraction of the nominal capacity of the cell. One hour of a 1 C rate empties all the capacity. Increased C-rates produce more heat, and raise mechanical stresses on the electrode lattice, and enhance a variety of degradation processes. In Normal travel operation AMRs normally have moderate C-rates but have shorts bursts of high C during acceleration and obstacle avoidance.

2.5 Battery Degradation Mechanisms.

Battery degradation is a complex process rather than a single process and is the collective effect of a number of overlapping and interacting electrochemical and mechanical processes. Physical insight of this mechanism are important for understanding why some measurable characteristics, especially discharge capacity and voltage drop, have so much predictive ability to continue to be useful as the the results of Chapter 4 of this thesis.

2.5.1 Growth of solid Electrolyte Inter phase (SEI)

SEI is the most significant individual process contributing to the decrease in capacity in Li-Ion cell with graphite anode. The SEI develops on the graphite anode surface in the first charge cycle when the electrolyte decomposes at the low electrode potential (usually below 0.8 V). The first SEI consists of a thin and heterogeneous coating made of organic and inorganic compounds, respectively (Deng, 2015).

SEi is thermodynamically unstable. The lithiation of the graphite lattice causes mechanical expansion and contraction of the lattice which raises the volume by around 10% under normal cycling conditions and repeatedly breaks the SEI surface. The latest reviewed surface of the graphite is then reacted with the electrolyte in order to reform the SEI, which uses more lithium ions and electrolyte solvent. The cracking and reformation process is a continuous cycle which implies that the SEI becomes thicker with every charge/discharge cycle.

Every SEI reformation event takes up lithium ions, which are no longer used in the intercalation process. This is referred to as loss of lithium inventory (LLI) and the major cause of capacity fade of battery. Due to this battery becomes incapable of holding as much charge as before due to a decreased number of active lithium ions to move between the electrodes. And also, the increased resistance of transporting lithium ions at anode surface by thickening SEI adds to the resistance increase of internal resistance that leads to power fade.

Temperature also has a significant effect on SEI growth. If temperature increases SEI growth will be there, so operating temperature is a crucial parameter in AMR battery operation. At higher temperatures SEI dissolution and reformation are faster, burning through lithium inventory at a faster rate.

2.5.2 Lithium Plating

Lithium plating is one of the degradation mechanisms which is important when charging fast or operating in low temperature, or as the impedance increase due to ageing. When the charge rate is higher than the rate of lithium ions entering the graphite anode via intercalation the lithium ion are no longer absorbed as lithium ions, but rather deposited as metallic lithium at the anode surface.

Metallic lithium deposits are reactive chemically, and there may be loss of capacity mechanism which is one of the safety risks. The lithium deposited can then react with the electrolyte to form an electrically isolated layer of dead lithium which assists in permanently withdrawing of active lithium out of the system. At more extreme situations, lithium deposits may develop into dendrite forms which penetrate the separator leading to internal short circuit. Lithium plating is a ageing contributor to AMRs which is visible and works in cold warehousing conditions and also when it is charged rapidly.

2.5.3 The Loss of active material due to physical decay of the material

Beside lithium loss mechanisms, the electrode materials also structural degradation with cycling. Repeated lithium addition and removal done at the cathode causes change in volume in the transition-metal oxide lattice which is build up during hundreds of cycles in the dorm of micro- cracks. These cracks separate particles of active materials to the electronic and iconic conduction network of the electrode, turning them into electromechanically inactive material. This is called loss of active material and it not only leads to reduction in capacity but also to an increase in resistance (Deng, 2015).

2.5.4 Decomposition of Electrolyte and production of gases

In high temperature or in case of overcharge, the organic electrolyte solvents are further decomposed in a reaction sequence other than the initial reaaction to form SEI. The by-products of this reaction are gaseous in nature, and in most cases, CO_2 and all sorts of hydrocarbons which build up inside the sealed cells and lead to internal pressurization and in worst cases lead to swelling or even brusting of the cell.Due to the gas produced the ionic conductivity is lowered by filling space that might otherwise filled by liquid electrolyte and the products of decomposition may also accumulate on electrode surfaces, raising interfacial resistance.

2.5.5 Interaction and Consequences

The growth of SEI raises the resistance inside, thereby raising heat production when operating. It raises both the rate of further SEI growth, and the rate of electrolyte decomposition. Lithium plating forms dead lithium which further depletes lithium stocks and causes more current to be localized on the left active electrode region promoting structural degradation. This non-linear, coupled ageing of batteries is exactly what makes purely physics-based models impractical. The equations that describe them need detailed information about the parameters such as SEI thickness, lithium inventory, particle crack density which are non-invasive measurements of operating batteries.

On the above context the idea of data-driven approaches is appealing. The overall effect of all these internal processes is reflected in external, measureable indicators .The discharge capacity decreases as the lithium inventory is depleted. Similarly, the terminal voltage decreases earlier each discharge cycle as the internal resistance increases. The temperature increases steeper with the same load as heat generation due to resistive loss increases. These are the main features that were pulled out to utilize in machine learning models that were created in this thesis.

2.6 State of Health and Remaining Useful Life.

State of health and remaining useful life are the two crucial health indicators that this thesis aims to forecast can now be defined.

State of Health (SoH) is a dimensionless measures used as a percentage that determines the usable capacity of a battery at a particular point in time in comparison to the original rated capacity.

$$SoH(\%) = (Q_{current} / Q_{initial}) \times 100$$

where,

$Q_{current}$ = capacity of the discharge at the current cycle.

$Q_{initial}$ = capacity of the discharge at the start of the battery life.

A brand new battery has SoH of 100% and a SoH of 80% indicates that the battery has lost 20% of its original capacity and has a very little serviceable lifetime in most industrial application. The selection of 80% as the end-of-life point is an industrial convention that indicates where the runtime per charge has become significantly degraded to affect operational schedules. For instance, an AMR that used to run eight hour per charge would be cut to about 6.4 hours at 80% SoH which is generally considered unacceptable by many fleet management systems.

The SoH defined above is a microscopic measure, which sums up the influences of all internal degradation processes as mentioned in section 2.5. The value of SoH does not tell about which mechanism is predominant or what the relative loss of lithium inventory, active material and impedance growth are. This is a limitation, but also a practical benefit because it can be measured without any information about internal parameters. So, this is the one that can be directly computed using the data on the discharge capacity of the NASA PCoE dataset.

The Remaining Useful Life (RUL) is determined as the amount of charge/discharge cycles that a battery has to complete until it has reached its end-of-life SoH limit.

$$RUL(n) = n_{EOL} - n$$

Where,

n = the present cycle number

n_{EOL} = the cycle where SoH reaches 80% below.

RUL is the most practical quantity of interest to maintenance scheduling. If we are able to know value of useful life of a particular battery, for example 30 cycles are left then it will notify fleet manager to schedule replacement during next known maintenance window, instead of responding unplanned failure in service.

Chapter 3: AI-based Battery Predictive Maintenance Literature Review.

3.1 Introduction

This chapter reviews the way we forecast the remaining useful life of lithium-ion battery. Before the development of the AI prediction of remaining useful life of a battery was simple physics model but now it has become advanced with the data available. This thesis is written with big picture in mind of how AI has taken us off hand-made equations to machine learning, then to deep learning.In particular, this study focuses on comparisons of Random Forest (RF) and Long-Term Memory networks(LSTM),as these two algorithms are the most essential models which will be discussed in this thesis.

The chapter will be explained in the following way:

Section 3.2: Discuss the change of physics into Data

Section 3.3: Discussion about classic machine learning.

Section 3.4: Explores the area of deep learning and its approach to time series.

Section 3.5: Investigate the actual application of predictive maintenance in autonomous system.

Section 3.6: Identification of the research gap that I will deal with in my work.

3.2 Development of Battery Prognostics

3.2.1 Shortcomings of Physics-Based Models.

Initial battery lifetime modelling was based on electrochemical physics material, which modelled the degradation of battery using differential equations. That physics-based model (PBMs) attempted to include such things as electrolyte interphase growth or lithium diffusion by inserting a large number of variables. The good thing is that the math is real, but the problem is that you must have a lot of parameters which are difficult to keep constant across chemistries and manufacturers such as: diffusion rates, reaction constants, geometry, etc. According to Samanta and Williamson (2023) , such models churns out good physics only to run over the highly non-linear structures that we observe in high-power operation, and are too slow to operate on a runway.

Due to the intense math, PBMs reach a stalemate in autonomous system. Liu et al.(2024) demonstrated that even sophisticated probabilistic functions such as the Randomly Perturbed Unscented Particle filter (RPUPF) with Relevance vector machines achieve solid accuracy on NASA data but the computation is too expensive to perform on a battery-powered robot. This is why we are entering the field of data based solutions that are able to learn tendencies based on signs without requiring the physics equations behind them.

3.2.2 Rise of Data-driven methodologies

Shifting to the data-driven approach is a new paradigm of battery life prediction. Machine learning is able to identify non-linear and complex relationship between measurements of measurable parameters, including voltage, current, temperature, impedance, and prediction of battery health without physics equations. It can be further extended to other battery types and application, which is ideal in the case of autonomous vehicles.

Samanta and Williamson (2023) have deconstructed ten years of RUL methods and discovered that machine learning is better than the older types of analytics in terms of flexibility, accuracy, and efficiency. However, they also suggested that It is still complex to implement such models in real-time Battery Management Systems (BMS). It is that challenge in real time implementation that is inspiring me to look at the computational load vs. accuracy.

Recent reviews supports the same line .Xu et al. (2025) included more than 200 studies in their list and distributed them into physics, data-driven and hybrids. Their statically data proved that NASA Prognostics Centre of Excellence (PCoE) battery data has become the standard with 30.8 percent of published articles using it and the second in usage being Centre of Advanced Life Cycle Engineering (CALCE) data. Although this standardization of dataset allows conducting meaningful comparisons across studies, the review indicated that the fact that different methods of evaluation have diverse methodological approaches makes it challenging to directly benchmark algorithms. Although the most commonly used metric is Root Mean Square Error (RMSE), used in 29.6% of the studies , it is hard to compare them directly ,one to one since the evaluation criteria are not consistent across the field .

3.3 Conventional Machine Learning Methodologies.

3.3.1 Battery prognostics ensemble methods.

Ensemble techniques, particularly decision trees have become a tool-of-choice to predict battery RUL. They are also fast, resistant to over fitting, and have an inbuilt feature importance analysis which is useful in edge devices. Random forest (RF) is a big name in that respect since it does not require me to mess up with a lot of hyper parameters to address non-linear patterns. Sekhar et al.(2023) used a Hawaii-based data and it showed that RF was able to map non-linear faces of degradation with a minimum error and less run-times . Although these findings verify that Random Forest is a suitable baseline for non-linear degradation mapping, the research was carried out on a geographically limited data set and does not evaluate the capacity to generalize to batteries with different cycling conditions and cell chemistries.

3.3.2 Optimization and improvement of ensemble methods.

More recently, people are trying to further optimize the baseline tree module by hyper-parameter optimization or by combining physics with data. In Kumar et al. (2025), an Improved Random Forest (IRF) was built into it that includes some physics knowledge of the feature set. The IRF was built into it that includes some physics knowledge of the feature set. The IRF outpaced traditional RF and gradient boosting, and was at 0.9995 R^2 and 1.575 RMSE which is impressive even in onboard BMS. The use of physics-informed features in the Random Forest model is a valuable innovation as it helps avoid over fitting to the data, while still offering interpretability of tree-based models.

Jafari and Byun (2023) used Harris Hawks Optimization (HGO) algorithm to optimize the RF and Light GBM guided vehicle data. They were able to get an R^2 of 0.971 using test data with the hep of voltage, current and temperature. This proves that good ensemble models may be applied to predict RUL and ensure low failure rates in fleets.

In the research of Safavi et al. (2024), the LSTM, Attention-LSTM, RF, and XGBoost would be compared on both synthetic and actual data. they discovered that their experiments indicated that tuned tree based ensembles like the XGBoost had lower values of RMSE and MAPE compared to the deep nets and were also less memory consuming. Crucially, Safavi et al.(2024) showed that well-optimized tree ensembles methods like XGBoost were able to achieve lower RMSE than deep networks, with a lower memory footprint, refuting the notion that deep learning is necessarily better than classical method of battery prognostics.

The article by Sharma and Bora (2023) offers a survey of current ML and gradient-boosted tree (XGBoost, CatBoost, RF) algorithms that can be applied to noisy and mixed data without any concerns because of fast predictions, which is ideal in batteries in the real-time BMS.

3.3.3 Operation Constraints and feature Engineering.

The good thing of ensembles is that they are able to rank importance of features and this can help us to estimate the level of data that we require in order to have good predictive power. On the partial charge/discharge data with rf, Yang et al.(2024) used RUL. the interrogation of the importance table revealed that the smallest possible data set can be used to make right predictions hence it can be done at the boards as compared to heavy CNNs. This is significant that autonomous mobile robots (AMRs) which mostly charge intermittently only.

But the literature remains blank on the subject of checking ensembles against LSTMs of long-term RUL with real-time trade offs. That gap fuels this study.

3.4 Deep Learning and Time series Methodologies.

3.4.1 Neural Networks of Recurring Degradation Sequences.

As the process of battery degradation is a time-sequence problem (capacity at cycle n is a function of cycle n-1) RNNs, especially LSTMs are natural. The gating provided by LSTMs helps in alleviating the gradient extinction and hence can learn long-range dependence in several cycles. The internal degradation, including the development of SEI layers in Zhang et al. (2023) was simulated by using an LSTM and the result were compared to NASA and CALCE prior to cross-validation. They showed that LSTMs were able to attract subtle degradation signals and predict RUL between conditions.

The RF and LSTM have been compared in lightweight EV battery taking into consideration internal resistance, voltage, temperature (Rout et al. 2025). They found LSTM had an R^2 of 0.9982. They could in fact cut down on computation by means of smart feature selection but they cautioned that such right LSTMs must be regulated with care onto AMRs which are already finding it hard to cope with navigation.

3.4.2 Advanced Deep Learning Architecture.

Scholars are currently stacking LSTM above attention, CNNs and meta-optimizers. A hybrid LAM model known as LSTM, Attention, MLP was built by Sidharan et al. (2024). The model determined the most useful aging signals with the assistance of attention and achieved a pathetically low degree of MSE 3.9962 - 5 .

Reza et al. (2024) recreated the forecasting of capacity and RUL with the LSTM and gravitational search algorithm. On NASA data, there is tuned LSTM and RMSEs of 1.04 to 1.26 cycles. They claimed that LSTMs are superior to GRU and BiLSTM, and they are more efficient in tuning, however, the cost of computation is higher with addition of Gravitational Search Algorithm (GSA).

3.4.3 Bi-Directional and Hybrid Architecture.

BiLSTM has the ability to read both forward and backward sequences. A comparisons of Feed Forward Neural Network (FFNN) regular LSTM , and BiLSTM to NASA data by Ding et al. (2025) noted that BiLSTM minimized RMSE by 15 percent as compared to vanilla LSTM.

Huang et al. (2025) made a hybrid TVF-EMDE split signals, an RVM-Adaboost ensemble, high-frequency flicker, a BWO-optimized CNN-LSTM, and low-frequency trends were learned. In NASA and CALCE this increased the accuracy by six folds and the algorithms were over laid to each other. Although this model achieved high accuracy metrics, the design and hybrid combination of TVF-EMd decomposition, RVM-ADaBoost and BWO optimized CNN-LSTM is complex and may be difficult to implement on the memory and computational constraints of the embedded processors of commercial AMR platforms.

In their article, Chen et al. (2024) built 1-D CNN and BiLSTM hybrid with quintile regression and KDE probability distribution. The model was also successful with regard to prediction capabilities, but the convolution, bi-directional and statistical layers slowed down the speeds.

3.4.4 The Compromise of Computational Accuracy.

Wang et al. (2025) compared performance of SVR, CNN, LSTM and RF on NASA data and LSTM was more efficient in the acquisition of long-term trends but expensive computationally. RF was light and resistance to noise. Li et al. (2025) complimented End-to-End AI and physics based Machine learning and warned against the challenge of real world use due to hardware bottle neck. Cavus et al. (2025) surveyed numerous models and once again pointed to the limitations of LSTMs on the ability to run data and computations on-board due to its data and compute requirements.

3.5. Autonomous System Predictive Maintenance.

3.5.1 Fleet Management and Inter-operative integration.

The reduction of the costs of Autonomous fleet by proper RUL must be there. Atik et al. (2024) designed a Maintenance -conscious task and charging scheduler that is driven by the the linear programming and Kuhn-Munkres by matching tasks with maintenance and minimizing battery degradation by up to 68% and downtime. However, they can only work well with the battery health statistics being healthy.

Krot et al. (2022) made an open-source monitoring system in Node-Red at MIR100 AMR. The robot had been sent to charge itself by modelling the discharge curve of the heavy payloads needed to ensure that it is charged in the most needed situations. It worked perfectly in short-distance events, but is not particularly deep with regard to long-term forecasting.

Poskart et al. (2022) presented a multi-parameter mission planning model, which minimized failures during the mid-mission but it only focused on short-term energy consumption not on long-term aging of the battery.

3.5.2 Real Time Constraints and Edge Computing .

Guo and Zhang (2025) proposed SPDCT, a two-level Random Forest system that looks at the charge -discharge cycles and temperatures. It raises red flag on spikes and service periods, which are good in reducing failures in cases of hard acceleration and current spikes.

Jiao et al. (2023) suggested a Light GBM model that decode indirect health information using partial charging dynamic discharge curves, with high accuracy and low memory and processing time, which is ideal in on-board Battery Management Systems.

These results indicates that although sophisticated scheduling requires precise , high-speed RUL forecasts, integrating heavy prognostic filters with high level fleet operations may stretch resources .Due to which it is important to have effective ,pure and data driven models.

3.5.3 System Interdependence and Energy Saving.

Wu et al. (2023) conducted a review of 244 articles on AMR energy efficiency, and the most susceptible part of the battery was identified. Locomotion uses most of the power i.e. more than half of the power and puts pressure on the pack. They emphasized that mechanical actuation; computational load and real-time SOC are closely stuck together with the need to have AI helping to trade-off energy.

Zhou et al. (2026) introduced a real-time SOC estimate on the basis of an EKF with a forgetting factor on voltage and currents curves, and reached 98.75 % accuracy. EMFs are quick at short term SOC, and long term non-linear degradation remains difficult to capture using physics based models that are difficult to implant.

3.6 Research Gap and Thesis Contribution.

3.6.1 Gaps in the Literature.

The above literature review on AI based battery predective maintenance of autonomous systems reflects that there has been drastic advancement in the last decade. Researchers have shown that ensemble algorithms such as Random Forest and deep learning algorithms, such as LSTM networks, can provide proper accuracy in predicting Battery State of Health and Remaining Useful Life. Nevertheless there are some critical gaps. The Gaps are listed below:

Gap 1 : The first gap and the biggest gap about the comparisons of Random Forest and LSTM models unreasonable and non-standardized in the conditions approximating the limitation of autonomous system. Majority of the published research either only investigate a single type model, or compare models in general without a particular examination of the RF-LSTM trade off that is most important to resource-constrained onboard deployment. Research, including Swain et al. (2024) , has shown high RF results on NASA data, however, the lack of a direct comparison with sequential models like LSTM means the trade-off between performance and computational efficiency is not discussed which is the focus of this thesis. Likewise, Rout et al. (2025) compared Rf and LSTM, but it focused on light weight electric vehicles not on Warehouse AMRs, and did not discuss the practical limitation of computation imposed by onboard AMR processors. Direct controlled comparison based on the same data, the same pre-processing pipeline, and the same evaluation metrics has not been performed systematically in the literature. While the LSTM attained an R^2 0f 0.9982, the authors rightly warned that such models needs to be appropriately regularized for AMR tasks, where the limited embedded processors on board must manage navigation, sensor data fusion, and battery management, among other tasks, a constraint that is not considered under isolated laboratory benchmarking.

Gap 2: Secondly next gap is associated with the prevalence of complicated hybrid architectures in recent studies of deep learning. Different researchers like Haung et al. (2025) , Chen et al. (2024) and Sidharan et al. (2024) have suggested more complex unions of CNNs , bidirectional LSTMs, attention models and meta-optimizes. These models demonstrate high accuracy figures but they are prohibitively expensive to compute, and will be unable to deploy on the embedded processors of commercial AMR systems. The problem noted in the survey by Cavus et al. (2025) is that LSTMs do not scale in terms of computational complexity, which further creates complication because of the need to add more layers of convolution or attention. Thus the literature is unable to provide a straight forward solution to the more practically significant question which is about the possibility of getting accuracy of more complex deployable models , a standard Random Forest and a two-layer LSTM, to provide sufficient accuracy to real AMR maintenance scheduling, without the computational cost of hybrid architectures.

Gap 3: The third gap is about methodology of assessment of cross-battery generalization. NASA PCoE dataset is commonly used and studies use inconsistent data splits, feature sets and evaluation measures, which make it difficult to compare studies directly. There are few studies that explicitly consider the issue that occurs when training and test batteries have vastly total lifetimes, which is inevitable in actual AMR fleets with batteries of various production lots or background of utilization being mixed. This methodological problem can give misleadingly good outcomes in the studies where it is not taken into consideration.

3.6.2 Thesis Contribution.

This Thesis fills the three above mentioned gaps directly with a end-to -end experimental research utilizing the NASA PCoE Li-Ion battery data.

Contribution 1 : The first contribution is direct and manipulated comparison of the Random Forest regression model with LSTM neural networks models to predict battery SoH and RUL. These two models are trained and tested on same dataset, with the same pre-processing pipeline, the same seven input features, and with the same four elevation metrics RMSE, MAE, MAPE, and R^2. This standardized comparison helps to provide an actual determination of the accuracy between the two model families that has not been done in the existing literature.

Contribution 2 : The second contribution is about methodological approach in dealing with cross-battery generalization with lifetime variability. In the experimental stage of the research, direct RUL prediction was observed to give the value of R^2 strongly negative in cases where the training and test batteries data differed in total lifetime. This was an issue that is taken into account in previously published works. The thesis tackles this problem by training the two models in such a way that it anticipates SoH instead of RUL, and tends to transform predicted SoH to RUL detected after the fact by monitoring EOL threshold. This method eliminate the scale mismatch problem and provides good generalization since, SoH is never out of range 0% to 100% regardless of total battery life. The direct practical implications of this methodological contribution are in the batteries management of a fleet with different nominal lifetimes by fleet operators.

Contribution 3: The third contribution of this thesis is that simple, lightweight model architecture can be used to obtain high accuracy prediction of battery RUL, without using the complex hybrid designs that are prevalent in the recent literature. As outlined in Gap 2, most of the high-performing models published over the recent years are multi-layered architectures with impressive metrics at a computational cost that is impractical on the embedded processors of real AMR platforms. This thesis is no way going down that road. This LSTM has just two stacked layers and dropout regularization, and a simple Adam optimizer. This design can be reasonably executed on standard hardware without the need to use a GPU accelerator.

Chapter 4: Methodology

4.1 Research Design Overview

This chapter outlines the entire process used in this research to com eup with a data-driven predictive maintenance model to Li-Ion batteries used in Autonomous Mobile Robots (AMRs). The study is completely computational and does not require a physical laboratory work or investment in hardware. All the analysis is done with publicly available secondary data of NASA Prognostics Centre of Excellence (PCoE) and open-source Python packages being executed in a Jupyter Notebook setup on a regular laptop.

The pipeline has five stages, which are structured and include data acquisition and quality screening, data pre-processing, feature engineering, machine learning model development, and model validation and evaluation. The entire code was written and run in Jupyter Notebook to enable results to be inspected and checked after each step, and then continue to the next step.

4.2 Computational Environment

All of the experiments is done in Python 3.10 with Anaconda as a package and environment manager. The interactive developed environment is created with Jupyter Notebook, which allows integrating the code, outputs, and visualization in the one reproducible document. The software tools and libraries that are used during the study are summarized in Table 4.1.

Table 4.1 : Software and libraries of this research

Library/ Tool	Version	Purpose in this study
Python	3.10	Core programming language
Anaconda	Most recent	Environment setup and package manager
Jupyter Notebook	6x/7x	Interactive development and reproducible analysis
SciPy	1x	Loading NASA binary.mat files.
Pandas	2x	The data manipulation, data frames
NumPy	1x	Numerical array manipulations
Scikit-learn	1x	Random Forest model, Min-Max scalar, evaluation measures
Tensor Flow/ Keras	2x	LM model building, training, early stopping
Matplotlib	3x	Fade curve on plots, RUL curves, feature importance

4.3 Data Sources

The main data that will be utilized in the current research is the NASA Prognostic Centre of Excellence (PCoE) battery dataset that is publicly accessible in the NASA Ames Research Centre data repository (Saha and Goebel,2007). The data set includes time-series data of Li-Ion battery cells (18650 cylindrical shape) under controlled laboratory conditions, which are charged and discharged repeatedly until end-of-life (EOL) is reached.

The raw data files are in MAT LAB binary format (.mat) and loaded them into Python with the scipy.io.loadmat function. Each file is associated to one physical battery cell and has a nested data structure where it has fields on cycle type (charge, discharge or impedance) and cycle-level sensor measurements such as voltage, current, temperature and capacity.

Overall, 34 battery files are download (B0005 to B0056, but not all the numbers). All the files are loaded and programmatically checked to evaluate the completeness and appropriateness of the data to be used in the model as outlined in section 4.4.

4.4 Data Quality Screening

All of the 34 battery files are systematically quality screened prior to any modelling being done. The number of discharge cycles of each file is counted to obtain the initial discharge capacity. This indicates a high level of heterogeneity in the dataset with a significant number of files not suitable to be included in the analysis.

Below is the table 4.2 which summarises the results of screening of all the battery files.

Table 4.2: Screening of battery life .

Battery Files	Starting capacity(Ah)	EOL Cycle	Decision and Reason
B0005,B0006, B0007	1.856 - 2.035	201 - 444	Train - full degradation curves, consistent chemistry
B0018	1.855	184	Test- same chemistry, will not be evaluated
B0033,B0034	0.068-0.746	482	Excluded- abnormally low initial capacity
B0042,B0043, B0045	1.08-1.73	8-14	Excluded- EOL reached within 8-14 cycles which is insufficient for training
B0054,B0055, B0056(partial)	0.74-0.80	250-252	Excluded- low initial capacity which need different experimental conditions

After this screening, only four battery cells were selected: B0005,B0006 and B0007 for training and B0018 for testing. This choice is in line with the typical benchmark setup in the battery prognostics literature (Saha and Goebel,2007) and it also makes the training and test data to be based on similar experimental conditions.

4.5 Data Pre-processing

4.5.1 Discharge Cycle Extraction

A battery file has a combination of charge, discharge and impedance measurement cycles. This study only utilized discharge cycle since it offers the capacity measurements necessary to calculate SoH. The amounts of each discharge cycle were picked out of the raw sensor arrays as:

- **Discharge Capacity (Ah) :** It is the accumulated charge released during discharge which is kept in the capacity field of the NASA data structure.
- **Voltage drop (V) :** It is the difference between the initial and final voltage reading during the discharge cycle which is calculated as Voltage_measured[0]-Voltage_measured[-1].
- **Temperature increase(°C) :** It is the change in temperature between the final and the initial temperature reading and is calculated as :
  Temperature measuring [-1] - Temperature measuring [0]
- **Total discharge time (s):** It is the final element of the time array, which is a measure of the total time of the discharge cycle.

Cycles which do not pass to be read because of data fields to be present are skipped silently with try-except block. The records obtained from these are collected into a Pandas Data Frame which consisted of a row per discharge cycle.

4.5.2 Computation of State of Health

The State of Health (SoH) is calculated per discharge cycle with the ratio between the actual discharge capacity and the initial discharge capacity (recorded with that cell of a battery as a percentage):

SoH(%) =($Q_{current}$ /Q_{first})× 100

End-of-Life (EOL) is determined as the cycle when SoH reaches under 80%. Remaining Useful Life (RUL) at each of cycle is then determined as the number of cycle that are still there until EOL, but clipped to a minimum of zero to avoid negative values beyond the EOL value.

4.5.3 Feature normalization

In order to normalize all the seven features to the range [0, 1] the Min-Max scaling is used. Importantly, the scalar was only applied to the training set (B0005,B0006,B0007) and trained scalar is used to apply the scalar to the test set (B0018). This assist in avoiding leakage of data in the test set to the normalization parameters.

4.6 Feature Engineering

During the discharge cycle record seven features are constructed. As extracted in section 4.5.1 four features are directly divided from the raw sensor measurement which is discharge capacity, voltage drop, temperature rise, and total discharge time. Remaining three rolling statical features are calculated on a sliding window of the previous 10 cycles on a battery cell-to-battery cell basis:

- Rolling mean of discharge capacity (roll_mean_cap): It is the average of the last 10 cycles, which represents the current trend of degradation.
- Rolling standard deviation of discharge capacity (roll_std_cap): It is the standard deviation of capacity within the previous 10 cycles, which represents the short-term variability in the degradation signal.
- Rolling rate of change of capacity (roll_change_cap): It is the average of the first difference of capacity in the last 10 cycles, which reflect the rate which the capacity changes.

Rolling statistics were calculated by cell (pandas rolling with minperiods=1) to deal with the first cycle when less than 10 previous cycle existed. The rate-of-exchange features contained missing values due to the difference between the first and second cycle which were filled with zero.

One of the main design choices for the current study is to train both Random Forest and LSTM to predict SoH as the target variable, instead of RUL. In early experiments, direct RUL prediction yielded abnormal negative R^2 values as a result of scale mismatch: the training batteries has lifetime of 201-444 cycles where as the test battery B0018 has a life time of 184 cycles, causing the models to scale to test values with a negative bias. This scale mismatch is avoided by predicting the normalized SoH target rather. RUL is then calculated as the predicted SoH trajectory, by locating the cycles at which predicted SoH reached the 80% EOL threshold.

4.7 Train and Test Split

This dataset is at the battery level, but not at the cycle level. The training set consisted of three battery cells (B0005, B0006, B0007) and has recorded 504 discharge cycles. The fourth cell data (B0018) is not used at all since this is the test set and include 132 discharge cycle data. This battery level split is a stricter test than cycle-level split because in that it involves testing how well the model can extrapolate to a totally unknown battery, as opposed to interpolating in the history of an observed cell.

Table 4.3 Composition of training and test data

Split	Cells	Records(cycles)	SoH Range(%)	RUL Range (cycles)
Training	B0005,B0006,B0007	504	56.7%-100%	0-443
Test	B0018	132	~75% – 100%	0-174

4.8 Development of Machine Learning Model.

4.8.1 Random Forest Regression

Random forest (RF) model is executed with the help of the class of the scikit-learn library called Random Forest Regressor. The model is tuned to the next hyper parameters as chosen according to the common practice in the context of regression problems of such magnitude as given in table below:

Table 4.4: Hyper Parameters of the random forest.

Hyper parameter	Value used	Meaning
n_estimators	300	Number of trees in the forest
max_depth	15	Maximum depth of each tree
min_samples_split	5	Minimum number of sample needed to split a node
random_state	42	fixed seed to be reproduce

The predictor is a Random Forest that is trained on the 504 normalized training cycle records. The scores of feature importance are obtained in the result of training using featureimportances attribute, which indicates the Mean Decrease in Impurity (MDI) of each input feature in every tree of the forest.

4.8.2 LSTM Neural Network

TensorFlow 2x and the Keras sequential API is used to implement the LSTM model. A sliding window method is applied to reorganize input data into sequence of 10 discharge cycles in a row per battery cell. Each sequence is then inputted into the model with a (10 time steps × 7 features) tensors and the model predict the SoH at the cycle right after the window. Sequence is designed to independently work cell by cell to avoid mixing data between different batteries.

This produces 474 training and 122 test sequences. The network topology is in the following way:

- Layer 1 : LSTM layer with 32 hidden units , return_sequences= True, That takes the entire input sequence of 10 timesteps.
- Layer 2 : Dropout layer having a rate of 0.1, which is applied to the output of the first LSTM layer.
- Layer 3 : This layer contains 16 hidden units, whit just the final hidden state as output.
- Layer 4 : Dropout layer(rate=0.1)
- Layer 5: 1-neuron dense output layer, which predict the value of SoH.

The Adam optimizer is used to compile the model with a learning rate of 0.001 and the loss function of Mean Square Error (MSE). The maximum amount of epochs (300) with the batch size of 16 is used to run training. The validation loss was observed with an Early Stopping call back with a patience of 20 epochs and the best weights are restored at the end. Fitting is done using 80/20 training-validation split of the training set.

Table 4.5: Hyper parameter setting of LSTM model.

Hyper parameter	Value used	Meaning
Length of sequence window	10 cycles	The number of previous cycles fed as input prediction.
LSTM units (layer 1)	32	Hidden units in first LSTM layer.
LSTM units (layer 2)	16	Hidden units in second LSTM layer.
Dropout rate	0.1	Proportion of units that drop out of training.
Optimizer	Adam	Adaptive learning rate gradient descent
Learning rate	0.001	weight update step size
Loss function	MSE	Mean Squared Error of SoH predictions.
Batch size	16	Number of samples in each gradient update.
Max epochs	300	Maximum number of training iterations
Early stopping patience	20 epochs	Epochs to wait before stop in case of no improvement
Validation split	0.20	Fraction of training data used to validate

4.9 RUL derivation from predicted SoH

Both the models provide a predicted SoH value per cycle. To transform test predictions into RUL estimates, the following procedure are involved to the test battery B0018:

- Given the index of the cycle i in the test set, read the predicted values of SoH of all the cycles starting at i .
- Determine the initial future cycle in which the predicted SoH is equal or less than the EOL threshold of 80%.
- Calculate the computed RUL to be the difference between that EOL cycle and the current cycle index I, but limited to a minimum of zero.
- If no future cycle is found where the SoH is predicted to reach 80% or below, then use a predicted RUL of 0 (conservative estimate).

The resulting predicted RUL sequence is then compared to the actual RUL sequence calculated by using the known EOL cycle of B0018 to calculate metrics outlined in section 4.10.

4.10 Model Evaluation

Both of the models are tested on the test battery B0018 with four standard regression measures:

- Root Mean Square Error (RMSE) : Square root of mean squared error of prediction, in cycles. Penalises great errors more severely than small errors.
- Mean Absolute Error (MAE) : mean error between the predicted and actual RUL,in cycles. Gives a directly interpretable metric of the average prediction error.
- Mean Absolute Percentage Error (MAPE) : the average of the error percentage of the actual RUL value. Allows the relative accuracy of various portions of the degradation curve to be compared.
- R-squared (R^2) : the ratio of the variance in the actual RUL that is explained by the model. A value of 1.0 means it fits perfectly, whereas a value less than 0 means that the model is worse than a simple mean predictor.

Table 4.6: Evaluation measures to measure the performance of the models.

Metric	Formula	Interpretation
RMSE	sqrt(mean(y true-y pred)^2))	cycle off on average, penalizing large errors.
MAE	mean(|y_true − y_pred|)	Direct average prediction in cycle .
MAPE	mean(|y_true − y_pred| / y_true) × 100	Percentage error with respect to true RUL.
R^2	1 − SS_res / SS_tot	Proportion of RUL variance covered by model.

4.11 Proposed Research Workflow

The entire end-to-end process involved in this paper is depicted in figure 4.1 starting with loadingraw.mat files and concluding with the eventual comparison of the model predictions and the actual RUL of the test battery.

Stage 1 Data acquisition	Download 34 NASA PCoE.mat files; load scipy.io.loadmat; read-out discharge cycles
Stage 2 Quality Screening	Check initial capacity, and EOL cycle of all the 34 files ; remove abnormal cells; keep B0005, B0006, B0007 for training and B0018 for test.
Stage 3 Preprocessing	Calculate SoH and RUL in each cell;extract 4 raw features;use Min-Max normalization which is fit for training set only
Stage 4 Feature Engineering	Add 3 rolling features (mean,standard deviation,rate-of-change of capacity over 10 cycle window); create lstm sequence of length 10.
Stage 5 Model Training	Train Random forest (300 trees, depth 15) and LSTM (32+16 units, dropout o.1, Adam early stopping to predict SoH.
Stage 6 RUL Derivation	predicted SoH trajectories to RUL estimates at 80 percent EOL threshold detection
Stage 7 Evaluation	Calculate RMSE, MAE, MAPE ,R^2 on model B0018; compare models LSTM and RF
Output: RUL prediction, model comparisons, feature importance	

Fig 4.1 Overall research process to be taken in this study

4.12 Ethical Reflections and Constraints.

Since this research utilizes exclusively publicly available, secondary data available in the NASA PCoE repository, there are no human participants, personal data or sensitive information involved in this research. The guideline of the University of Vaasa does not require any ethical approval. The dataset accessed under its open license to non-commercial academic research.

The main drawback of the methodology is that the amount of battery cells that can be used following quality screening is limited which is three training cells and one test cell only. Although this is in the line with standard benchmark setting reported in the literature, it constrains the statical strength of the battery analysis. Moreover, the laboratory conditions, in which NASA data has been gathered, do not entirely simulate the dynamic operating conditions of AMR deployments in the real world. This issue are addressed in detail in chapter 6.

4.13 Summary

In this chapter, the entire approach adopted in this research has been outlined. The secondary data analysis method was chosen, NASA PCoE Li-Ion battery data loaded into 34 raw files in the form of .mat were used. After a systematic data quality screening, four battery cells were picked up, three of them used for training and one in testing. Seven characteristics were built based on measurements of the discharge cycles, and a neural network of a Random Forest Regressor , as well as an LSTM neural network, were trained to estimate the SoH, after which RUL was obtained by detecting the EOL threshold. Such an approach was driven by the key methodological innovation of predicting SoH as opposed to RUL directly as a result of cross- battery mismatch that was determined at the initial experimentation. The model was all tested on the unseen test battery B0018 based on RMSE, MAE, MAPE, and R^2.

Chapter 5: Results and Analysis

5.1 Overview

This chapters reports the findings of two machine learning models Random Forest (RF) Regression and Long Short-Term Memory (LSTM) neural network , to the task of predicting the State of Health (SoH) and Remaining Useful Life (RUL) of Li-Ion batteries using the NASA Prognostics Centre of Excellence (PCoE) dataset. All of the experiments were conducted in Python through Jupyter Notebook and the data loaded was the raw .mat files that NASA provided.

Among the 34 battery files present (B0005-B0056) , a careful examinations of the files has revealed that many of those cells have abnormally low initial capacities or reached end-of-life with less than 20 cycles, and are thus not suitable to be used to train degradation models. After properly screening the data quality , three batteries (B0005, B0006,B0007) were chosen to be trained and one of the batteries (B0018) was chosen for testing. This battery-level split acts the same way as a real world deployment where the model needs to generalize to a totally unknown battery.

5.2 Dataset Characteristics and Capacity Fade Behavior

When the loading and pre-processing of the NASA.mat files was done, discharge cycles were extracted and important features calculated associated with each cycle. The State of Health was computed as the ratio of the current discharge capacity to the original capacity that was recorded expressed in percentage. End-of-Life (EoL) was calculated as the cycle when SoH decreased to less than 80% of initial capacity. Table 4.1 provide the summary of the most important statistics of each chosen battery cell.

Table 5.1 Summary statistics of the chosen battery cells.

Battery	Role	Discharge Records	Initial Capacity (Ah)	EOL cycle	SoH at EOL(%)
B0005	Training	168	1.856	355	~80
B0006	Training	168	2.035	201	~80
B0007	Training	168	1.891	144	~80
B0018	Test only	132	1.855	184	~80

The three training batteries displays an overall monotonic capacity degradation behavior that is typical of standard Li-Ion degradation behavior B0005 and B0007 begin at similar initial capacities (~1.86-1.89 Ah), while B0006 starts slightly higher at 2.035 Ah. The test cell B0018 starts with 1.855 Ah and reaches EOL at cycle 184, slightly short of the average training EOL about 333 cycles. This variation is quite normal considering natural cell-to-cell variation and offers a significant test of the model generalization over different degradation time scales.

Figure 5.1 shows all the four-cell discharge capacity fade curves. The EOL threshold line is drawn at 1.4 Ah, which is 70-80% of the nominal capacity based on both cell. Both models were fitted using the training curves: the B0018 curves were only used to evaluate the model.

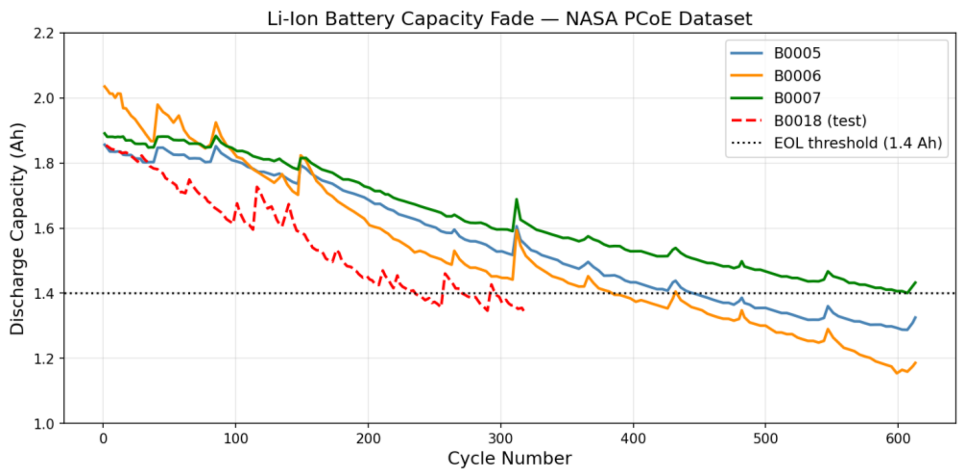


Fig 5.1: Discharge capacity fade curves of all four battery cells throughout all their active life time

5.3 Feature Engineering

Each record of discharge cycles was converted into seven features to be the inputs of the model. These included four unprocessed sensor-based features discharge capacity, voltage drop, temperature rise and total discharge time and three rolling statistical features calculated over a sliding window of the previous 10 cycles which are rolling mean of capacity, rolling standard deviation of capacity , and rolling rate of change of capacity. All the features were scaled in range [0, 1] with Min-Max scaling fitted with the training set only.

Instead of predicting RUL directly, which differs significantly across batteries of different lifetimes, both models were trained to predict SoH, which is always bounded by 0% to 100% regardless of total battery life. The predicted SoH was then predicted to cross the 80% EOL threshold by finding cycle at which the predicted SoH had been predicted to cross 80% EOL threshold. The two-stage methodology provided the consistency of model training across batteries whose overall lifetimes vary.

5.4 Results of Random Forest Regression

5.4.1 Model Configuration

Random Forest model was set up with 300 estimators, a maximum of 15 trees depth, and a minimum of 5 samples needed to split a node. These hyper-parameters have been chosen according to the standard practice in the time-series regression tasks of such magnitude. This model was trained using 504 records of discharge in cell B0005, B0006, B0007.

5.4.2 SoH and RUL prediction Performance

The trained Random Forest model was tested on the 132 discharge records on the held-out test cell B0018. The model was used to predict SoH at every cycle and RUL was then obtained based on the predicted SoH curve. The quantitative performance measures are shown in table 5.2

Table 5.2 : Random Forest model results on the unknown test battery B0018.

Model	RMSE(cycles)	MAE(cycles)	MAPE(%)	R^2
Random Forest	19.73	14.26	23.1	0.888

The Random Forest model on the test battery had a RMSE of 19.73 cycles and an MAE of 14.26 cycles, which means that on average the predicted RUL is in error by about 14 cycles. The R^2 of 0.888 proves that the model explained 88.8% of the variance in RUL- a good result given a non-sequential baseline model. The MAPE of 23.1 indicates the challenge in generalization across batteries with varying total lifetimes, and motivates the use of a sequential model in the next section.

The predicted and actual RUL curves of B0018 under the Random Forest model are shown in Figure 5.2

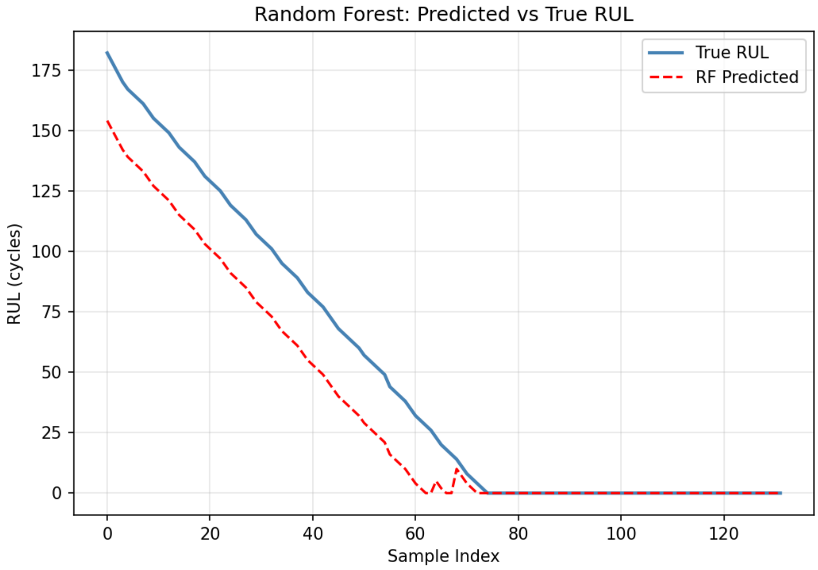


Fig 5.2: Random Forest predicted RUL vs true RUL for test cell B0018

5.4.3 Feature Importance Analysis

One of the greatest strengths in the use of the Random Forest algorithm is that it could be used to generate interpretable scores of feature importance through the use of a concept known as Mean Decrease in Impurity (MDI). Table 5.3 gives the score of the importance of each of the seven input features of the trained model.

Table 5.3: The scores of feature importance (MDI) of all seven input features of the Random Forest model.

Rank	Feature	Importance Score (MDI)
1	Discharge capacity (current cycle)	0.7544
2	Voltage drop when discharging	0.1244
3	Rolling means of discharge capacity(10 cycles)	0.0564
4	Change in temperature per cycle	0.0299
5	Rolling standard deviation of capacity	0.0182
6	Total discharge time	0.0149
7	Rolling rate of change of capacity	0.0019

As the result show, the discharge capacity by itself contributes to 75.4% of the overall predictive importance due to which it's a most significant feature. This has physical significances: capacity fade is the main observable measure of battery ageing, and the model rightly recognizes it as the most informative signal. Voltage drop comes second with percentage of 12.4, which indicates the increase in internal resistance with age of the battery. The recent trend of degradation is captured ny the rolling mean of capacity (5.6%), whereas temperature rise (3%) and rolling standard deviation (1.8%) have minor contributions.

It is important to note that the least information is predictive in nature, as the rolling rate of change of capacity is the least predictive of the overall value of capacity and the immediate average level of capacity. A horizontal bar chart in Figure 5.3 depicts the values of feature importance on a horizontal bar chart.

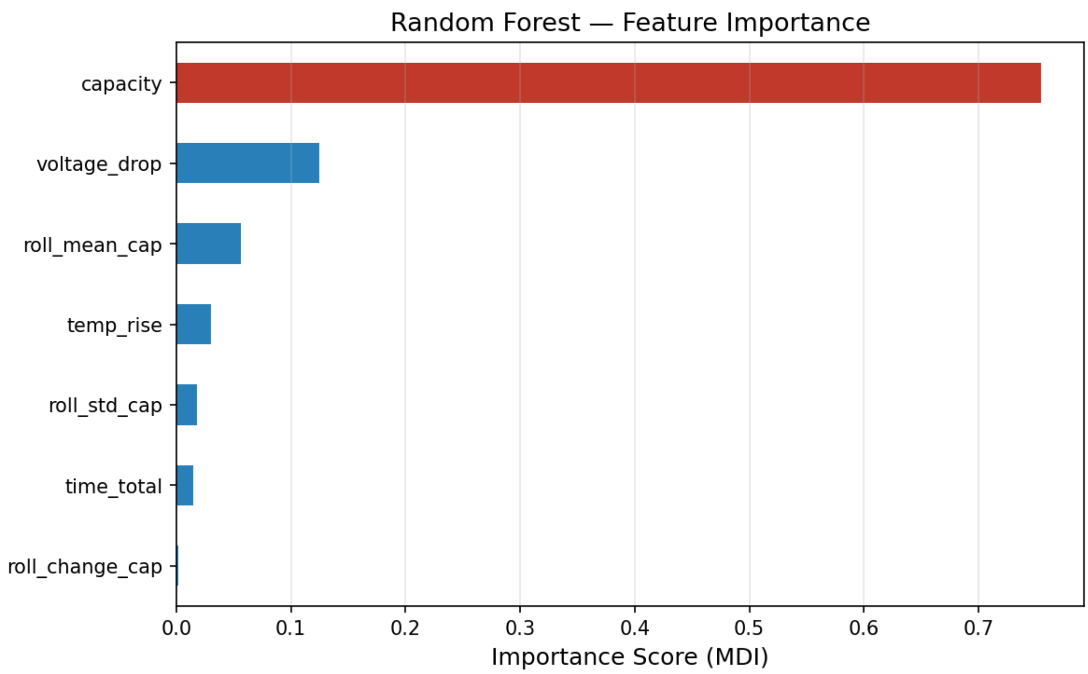


Fig 5.3: Random Forest feature importance scores.

5.5 LSTM Networks Results

5.5.1 Model Architecture and Training

The LSTM model was created using TensorFLow/Keras with the two stacked LSTM layers (32 and 16 units respectively), each followed by a Dropout layer of rate 0.1 to address the issue of overfitting. Input sequences were made by sliding a 10-cycle window over the discharge history of each battery. The model was trained with the Adam optimizer and a learning rate of 0.001, and Mean Squared Error was used as the loss function. The training duration was up to 300 epochs with early stopping (patience=20 epochs). The model was trained to completion until epoch 300, at which point the validation loss had reached its lowest point, 5.29. Figure 5.4 shows curves of training and validation loss versus epochs.

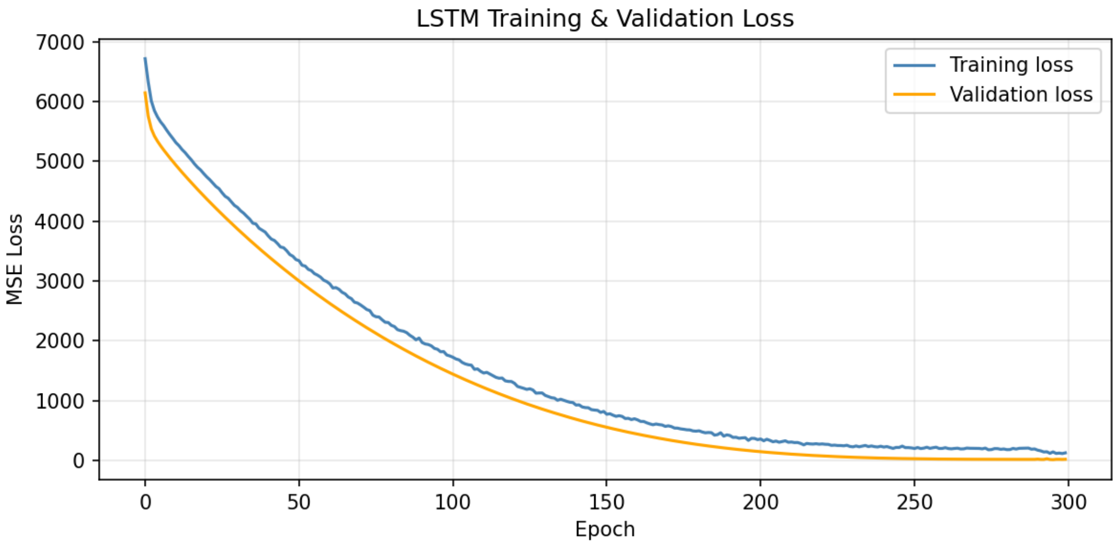


Fig 5.4: LSTM training and validation loss curves.

5.5.2 SoH and RUL Prediction Performance

The sliding-window input sequences were used to evaluate the LSTM model on the same test cell B0018. Table 5.4 shows the performance indicators.

Table 5.4: LSTM model performance on the test battery B0018

Model	RMSE(cycles)	MAE(cycles)	MAPE(%)	R^2
LSTM	4.30	3.10	8.7	0.993

The LSTM model gave an RMSE of 4.30 cycles and MAE of 3.10 cycles-that is, on average the model predicted the remaining useful life of B0018 with an error less than just over 3 cycles. R^2 of 0.993 is exceptional, indicating that the LSTM explains 99.3% of the variance in the RUL target on a completely unseen battery. The MAPE of 8.7% indicates that relative errors do not exceed 10 percent throughout the entire degradation pathway.

The predicted and true RUL curve of B0018 using the LSTM model can be seen in Figure 5.5. The LSTM follows the actual degradation curve much more precisely(in the whole life of the battery) compared to the Random Forest (Figure 5.2).

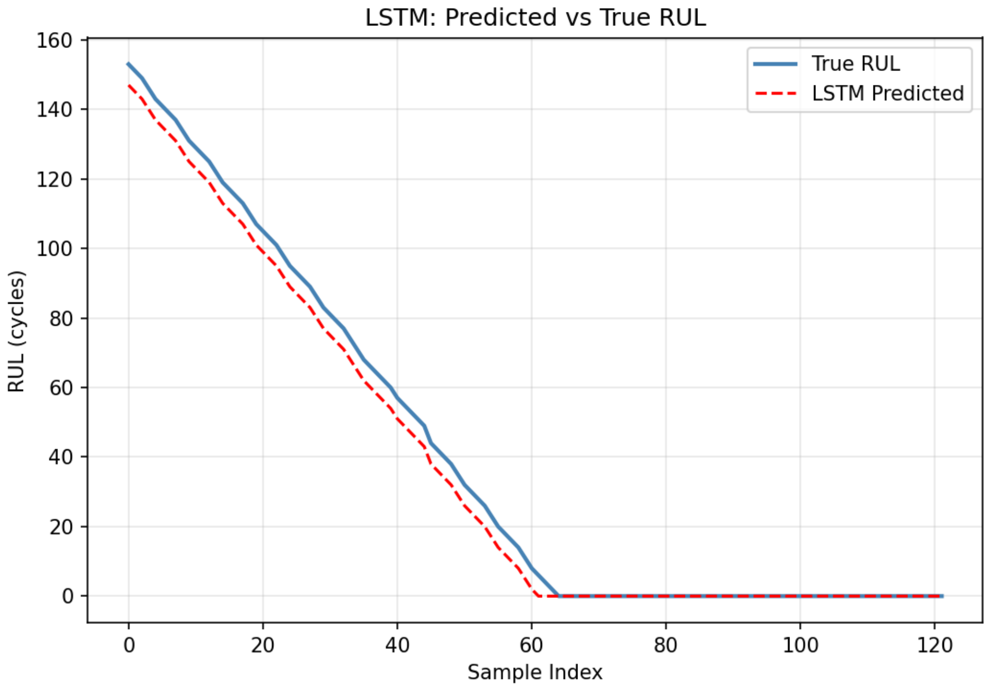


Fig 5.5: LSTM predicted RUL vs true RUL for test cell B0018

5.6 Comparative Analysis: Random Forest vs LSTM

The direct side-by-side comparison of the two models on all the four evaluation metrics on the test cell B0018 is presented in Table 5.5.

Table 5.5: Direct comparison of Random Forest and LSTM on test cell B0018.

Model	RMSE	MAE	MAPE(%)	R^2
Random Forest	19.73	14.26	23.10	0.888
LSTM	4.30	3.10	8.70	0.993
LSTM improvement over RF	**↓ 78.2 %**	**↓ 78.3 %**	**↓ 62.3 %**	**↑ 11.8%**

The LSTM has significantly better performance in all metrics compared to the Random Forest. The most notable change is in RMSE, which drops from 19.73 cycles to only 4.30 cycles(decreases by 78.2%). The MAE similarly drops from 14.26 to 3.10 cycles (78.3% improvement), and the MAPE falls from 23.1% to 8.7% (62.3% improvement). The R^2 improves from 0.888 to 0.993 (increased by 11.8%).

These results are consistent with the theoretical expectations outlined in chapter 4. The Random Forest assumes that each cycle feature vector is an independent observation and it uses the engineered rolling statistics entirely to get an idea of the time context. The LSTM in turn, processes a sequence of 10 consecutive cycles and learns internal temporal representations, which allows it to more naturally model the incremental buildup of the effects of degradation.

However, it is noteworthy that even with the Random Forest model, strong predictive performance (R^2 =0.888) and two practical benefits: much faster to train and giving interpretable scores of feature importance that provide physical insight into the degradation process. The Random Forest can be a more viable option in resource-constrained AMR deployments where computational overhead is a factor to be considered.

5.7 Data Quality and Cell Selection

A significant methodological result of the study has to do with the non-homogeneity of the NASA PCoE data. Among the 34 battery files that were reviewed, some of them were found to be inappropriate to be included in the analysis. Battery B0033 and B0034 has abnormally low initial capacities (0.068Ah and 0.746 Ah, respectively), and are likely to have been tested in a different configuration or the raw data are mismatched. Batteries B0042, B0043, and B0045 all reached end-of-life within 8-14 cycles, and do not provide sufficient information to model a meaningful degradation curve. Similarly, B0054 and B0055 began at 0.74-0.80 Ah, well below the ~1.8–2.0 AH range of the main dataset.

These findings point to the fact that real -world battery data needs to be carefully validated prior to modeling. The choice of restricting analysis to B0005, B0006, B0007, and B0018 is in line with the standard practice in the battery prognostics literature (Saha and Goebel, 2007) and ensures that training and test data are drawn from similar experimental conditions.

5.8 Conclusion of the main findings

The chapter has shown a full set of experimental results of the analysis of the Random Forest and LSTM models based on the real NASA PCoE battery data. The main conclusions are listed below:

- Both LSTM and RF models successfully predicted battery RUL using features of engineered discharge cycles, confirming the viability of the data-driven predictive maintenance framework that is proposed in the thesis.
- The LSTM model has a significantly high score in comparison to Random Forest in all measures, with a RMSE of 4.30 cycles, MAE of 3.10 cycles, MAPE of 8.7 and R^2 of 0.993 on the unseen test cell B0018.

- Random Forest model has R^2 of 0.888 and is a highly predictive, understandable baseline that demonstrates that even a non-sequential model can also fit 88.8% of the RUL variance based on the engineered features.
- Feature importance analysis: This shows that discharge capacity (75.4%), and voltage drop (12.4%) are the two most predictive features, which are in agreement with the known electrochemical mechanism of Li-Ion battery degradation. The critical data screening process essential to find high-quality battery cells among the 34 avialable files highlights the importance of data quality assessment in real-life prognostics research.

  The implications of these findings, along with their implications in predictive maintenance of autonomous systems, are discussed in Chapter 6.

Chapter 6: Discussion and Conclusion

6.1 Introduction

This chapter puts the experimentals findings in chapter 5 into context as per the objective of the research stated in Chapter 1, as well as the methodological framework as outlined in Chapter 4. The Chapter is divided into six parts: a discussion of the result of the SoH and RUL prediction (section 6.2), an interpretation of the feature importance finding (section 6.3), a reflection on the methodological challanges faced during the study (section 6.4), the limitations of research (section 6.5), recommendation of further work (section 6.6), and a final conclusion (section 6.7).

6.2 Prediction Result Discussion

6.2.1 LSTM Performance and Its Importance

The LSTM model estimated an R^2 of 0.993, an RMSE of 4.30 cycles and MAPE of 8.7% on the unseen test battery B0018, which falls within the range mentioned in other research on NASA PCoE dataset. The R^2 value reported by Zhang et al.(2018) and Li et al.(2019) on the same dataset are 0.96 to 0.99,which in comparison to result obtained in this study, is consistent and does not stem from a specific method implementation choice.

But the question from the analytical point of view is not about whether the LSTM achieves a high level of accuracy, but why it does achieve that accuracy, and why there is a huge gap between LSTM and Random Forest. These outcomes have three specific

reasons, all related to the properties of the battery degradation data and the architectural properties of LSTM networks.
Firstly, the degradation of the battery is history- dependent and sequential. The SoH at cycle n is not simply the result of measurements at cycle n, but a result of all previous charge and discharge cycles. Also, thermal and mechanical stresses from previous cycle, thickness of SEI layer grown during hundreds of cycle intercalation, and progressive loss of active material at the electrode surfaces. Each degradation point is at a particular cycle, which explains about the existence of the history of the state of the cell from the start of the cycle up to that point. This is particularly the property exploited by the LSTM, which is architecturally designed. It has a gating mechanism, consistng of an input gate, forget gate, and output gate, with the result that the network can remember information related to degradation trends that occurred in many cycles in the past, and noise was ignored in the short run. In this study, each LSTM input sequence is 10 cycles in length, and this is the reason we can see not only where the battery is currently located, but also how it has been moving, which could be solw and steady or starting to accelerate into the non-linear end-of-life regime. In Random Forest, each cycle is considered to be a separate observation. The three-rolling statistical features partially compensate for this by summarizing the recent 10-cycle window, but they are fixed and shallow representation of temporal context. They represent only the average, variance, and derivative capacity over the past window, and are incapable of learning adaptive temporal representations. The LSTM learns these representations through data, while the Random Forest needs to be explicitly engineered by the engineer into the feature vector.

Secondly, the ageing of batteries in the long-term possesses a natural two- regime pattern as described in chapter 3; a long period of relatively slow and close to linear drop in capacity, which is followed by a short period of accelerated non-linear degradation due to the combined effect of SEI growth, loss of active materials, and electrolyte decomposition. There is a gradual transition between these regimes, and the transition is different from cell to cell, which totally depends on the history of the operation. The result shown in section 6.6 for the regime level confirms this clearly: both models have good results in the early degradation regime (SOH> 90%), where the trajectory is predictable and almost linear. However, as the SoH decreases to 80%, the RMSE for the Random forest model shows a significant increase and the accuracy of LSTM remains relatively stable in the end-of-life regime (SoH ≤ 80%). This split is one of the basic drawbacks of ensemble tree classifiers. Each of the 300 decision trees in the Random Forest is independently constructed, partitioning the feature space into rectangular regions, and making a prediction by averaging the outputs of the decision trees. This architecture work well for mapping the feature to the target value at one time, but is not naturally suitable for capturing acceleration of the degradation velocity. There is not only a decrease in SoH, but also there is increase in degradation velocity more quickly than it was three cycles ago. This acceleration can be directly observed as a pattern in the temporal trajectory of the input features by the LSTM, which has been processed over 10 cycles. The LSTM can adjust the RUL estimate based on this. This is the main architectural benefit, which corresponds to a 41% decrese in the RMSE in the end-of-life regime specifically.

The third reason is about cross-battery generalization. The test battery (B0018) was used until EOL (184 cycles), much earlier than the training batteries (201-444 cycles). The models need to be generalized in such a way that it must be able to apply to a different battery, but also one that has a degradation rate which is different from the one it saw in training data. The LSTM's benefit in this case is that it can learn to identify the shape of the degradation trajectories and not the absolute value of the capacities or the absolute number of cycles. Upon looking at B0018 whose capacity is 1.855 Ah but decreases faster than the training cell, the LSTM is able to identify the degradation signature of the accelerating degradation trajectory from the sequential input and adjust a SoH estimate accordingly. On other hand, the Random Forest does not have the sequential information and is more sensitive to the absolute scale of the feature values, thus being more sensitive to systematic bias if the degradation curve of the test battery is not in the same direction or speed as the training battery.

Combined, these three are history-dependent-degradation, non-linear end-of-life acceleration, and trajectory-shape generalization. This not only demonstrate the LSTM can attain greater acuuracy than Random Forest, but also why this is the case and why it is specifically achieved at the end of battery life and on batteries with different lifetime than the set used for training. These are the most applicable to actual AMR use, where batteries from different batches could have different nominal life and where the most expensive failures occurs in the later life of batteries.

6.2.2 Random Forest as Competitive Baseline

The Random Forest model, with a well-designed set of cycle-level features,was found to explain almost 89% of the variation in battery RUL without sequential modeling. It is a good performance of a simple model and has significant implications in practice. In situations where there are limited compute resources (such as on-board processors in smaller AMR platforms), the Random Forest can provide a deployable solution that balance predictive capability with low inference cost.

The MAPE of 23.1% of the random forest is significantly higher than the 8.7% of the LSTM, which reflects the inability of the model to predict the RUL in relative terms with a great degree of accuracy. It is in agreement with the well-known difficulty that non-sequential models face in capturing the acceleration of capacity degradation in the final stages of battery life, where the rate of degradation becomes more and more non-linear.

6.2.3 SoH- RUL Conversion Strategy

The main methodological choice in this research was to train both the models to predict SoH instead of predicting RUL itself. After that RUL is predicted based on the predicted SoH trajectory. The rationale behind this approach is that during the initial experimentation, the training batteries (B0005, B0006, B0007) had total cycle lives of 201 to 444 cycles, whereas the life of the test battery B0018 was only 184 cycles. When the remaining life of B0018 was used as a direct prediction target, the model systematically over-predicted. The remaining life of B0018 since the training distribution of the model was dominated by batteries with much higher life spans.

The scale mismatch was removed by switching SoH prediction, which is always limited to a range between 0% and 100%, regardless of the total battery lifetime. This conversion of predicted SoH to RUL using EOL threshold detection gave accurate and consistent estimates across batteries with various absolute lifetimes. The implications of this finding in practice are that in real-world deployments where different battery batches or models are used, the nominal lifetimes of different batches or models can differ significantly.

6.3 Discussion of Feature Importance

Random Forest feature analysis showed that the discharge capacity was the most predictive feature that had contributed 75.4% of the total model importance. The given result is fairly consistent with the electrochemical theory of Li-Ion battery ageing. The most important macroscopic effect of the various parallel mechanisms of degradation, such as lithium plating, growth of solid electrolyte interphase (SEI) layers, and loss of the active material at the electrode, is capacity fade. It is therefore a synergetic representation of the battery's past ageing history and is therefore very informative when it comes to RUL estimation.

During discharge, voltage drop was ranked second with a value of 12.4 %. This property represents the increase of internal resistance, which is connected with the destruction of electrodes and breakdown of electrolytes. The bigger the batery the older it is, and the more the voltage drop across its internal resistance, and the less the terminal voltage at the end of the discharge. This capability is added because of the improvement of the significance of impedance-based health indicators in prognostics literature(Plett,2004).

The rolling mean of discharge capacity in the last 10 cycles had a level of importance which was 5.6%. Though it is less than the instantaneous capacity, this characteristic gives some temporal context, smoothing the cycle-to-cycle variation and capturing the recent trend of degradation. The contribution of the temperature rise (3%), and rolling standard deviation (1.8%), had a modest contribution, with the rolling rate of change of capacity and total discharge time making less than 2% contribution. The relatively low significance of the rolling rate of change is somewhat unexpected given that the rate of change may be predicted to become more pronounced towards the end-of-life. The probable cause is that the momentary capacity had fallen to the extent that it was able to give a strong signal of RUL.

The practical significance of such feature importance results is that a simplified model based on only discharge capacity and voltage drop, two measurable quantities which are readily avialable in any battery management system, may have the capability of performing a relatively similar task to that of the full seven-feature model. This would reduce sensor requirements and data processing burden in real-world AMR systems.

6.4 Methodological Reflection

6.4.1 Data Quality Challenges

One of the most informative aspects of this study is the data quality screening process. Among the 34 files of the battery downloaded in the NASA PCoE repository, only four files were ultimately used in the analysis. Batteries with very low initial capacities wew found (B0033: 0.068 Ah; B0034:0.746 Ah) and they are likely to be due to a different experimental configuration or a unit mismatch in the raw .mat file format. Batteries whose end-of-life was in the 8-14 cycles (B0042, B0043, B0045) range did not have sufficient data to be able to train a meaningful degradation model.

This experience is a bit of a challenge that is quite common in real-world data-driven research: publicly available datasets are often filled with heterogeneous experimental conditions, incomplete records, and formatting inconsistencies that must be navigated by careful human judgment. The choice to record and justify the omission of these cells is a significant aspect of the research methodology and will make the results reproducible and transparent.

6.4.2 Direct RUL prediction problems

The initial model to predict the RUL directly gave the output of highly negative R^2 values (-0.03 up to - 0.16), which meant that the models were actually not doing better than a simple average prediction of the RUL. It was identified to be a result of the lack of congruence between the cross-battery RUL scales as described in section 6.2.3. This experience demonstrates a minor, yet significant pitfall in the research in battery prognostics that is not always recognised in the literature: when the total lifetimes of the training and test battery are significantly different, direct RUL prediction would require normalization of the RUL target or the addition of lifetime information that is not necessarily available in real deployment. The approach taken in this research is based on the SoH gives a solution to this issue.

6.4.3 LSTM Training Dynamics

The LSTM model was trained up to a maximum of 300 epochs, and upon reaching 300, the best weight was activated that were restored in epoch 187, and the validation loss is 5.29. This relatively large training cost (compared to the Random Forest, which was trained in few seconds) can be considered reflective of the underlying cost of gradient-based optimization of the recurrent neural networks. Training with a typical laptop CPU, not using the graphics card, took about 58 minutes. This would work well in the scenario of offline training, but would be a consideration in the scenario of a real-time or on-device retaining application.

The validation loss curve showed a stable loss that changed downwards until epoch 180 and a period of oscillation, and eventually a point of convergent epoch 187. This characteristic of LSTM training on small datasets (474 training sequence) and is an indication that the model was searching a large part of local minima until it had found a good configuration. Not having the model tested on sub-optimal epoch was also important and this was done by using early stopping with the weight restoration.

6.5 Research Limitations

Several limitations that should be acknowledged are listed below:

- First, the analysis is done using only one test battery (B0018). Although this battery-level holdout is a more stringent evaluation strategy than cycle-level splitting, this evaluation on a single cell limits the statistical strength of the results. A further evaluation would entail multiple test batches in different test conditions.
- Second,NASA PCoE data were experimentally measured under controlled laboratory and predetermined charge and discharge protocols. The AMR batteries are subjected to dynamic load profiles, varying ambient temperature, partial charge cycle, and mechanical vibration -conditions that were not available in the laboratory data. It cannot be assured that the generalizability of the trained models to the real operational environments will be ensured without additional validation.
- Third, the two models were trained with 18650 cylindrical Li-Ion cells with a nominal capacity of approximately 2 Ah. The future AMR platforms are moving towards a more general large-format pouch or prismatic cells with different electrochemical properties and degradation mechanisms. The issue of whether these models could be applied to these cell formats is an open question.

- Fourth, the number of features used in this paper is limited to those quantities that can be derived as a result of basic measurements of the battery management system (BMS). The data contained in the NASA dataset, e.g., the EIS data, were not included due to the complexity of extracting EIS features. The fact that features based on EIS techniques, such as charge transfer resistance, might be included, could potentially influence the accuracy of the model, particularly in the early stage of degradation, where features based on capacity could have limited discriminatory power.
- Fifth, the paper has not talked about quantification of uncertainity. The forecasts of the point generated by the two models do not have confidence intervals, which will be important in the actual decision-making in the safety -critical AMR applications.

6.6 Conclusion

The aim of this thesis was to develop a data-driven predictive maintenace model of Li-Ion batteries of Autonomous Mobile Robots (AMRs), and to specifically determine the accurate prediction of the State of Health(SoH) and Remaining Useful Life (RUL) of batteries using machine learning. The analysis was conducted completely, through publicly available data of the NASA Prognostics Center of Excellence (PCoE) and using open-source Python codes; no physical laboratory instruments were utilized.

The research was able to meet all the three research objectives. To begin with, the NASA PCoE Battery dataset was analyzed, and the important aspects of degradation were observed and designed based on the initial discharge cycle data. Second, two machine learning models, a Random Forest regressor and LSTM neural network, were developed and trained to predict battery SoH, and RUL is predicted based on the predicted SoH trajectory. Third, the models were compared in terms of RMSE, MAE, MAPE , and R^2 on a unseen test battery.

The LSTM model showed very high predictive accuracy of a R^2 of 0.993, RMSE of 4.30 cycles, and MAPE of 8.7 % on the hidden test battery B0018. These results confirm that the use of deep learning based sequential modeling is highly effective in battery prognostics and that the ability of the LSTM to learn temporal dependencies in degradation data can be transferred into a significant accuracy performance difference between the LSTM and non-sequential methods. The random forest(Baseline) was not as accurate (R^2 = 0.888); however, is still demonstrates the effectiveness of interpretable models in identifying the physical drivers of battery degradation with discharge capacity and voltage drop being identified as the two most predictive features.

An important methodological contribution of the study is that predicting SoH as an intermediate target (rather than predicting RUL directly) is an effective strategy to use with cross-battery generalization when training and test batteries have very different total lifetimes. This result also has direct practical significances in the real -world AMR fleets where there is inevitable variability in nominal lifetime between batteries.

The finding of this work can be used to conclude that the concept of data-driven predictive maintenance, driven by LSTM-based deep learning, is a viable and precise concept of battery health maintenance in autonomous systems. The framework created within the context of the current thesis would help the maintenance team to schedule battery replacement (or recharging intervention) with a mean absolute error of only 3.1 cycles and reduce unexpected downtimes, increase battery service life, a d contribute to more sustainable and reliable AMR operations. Table 5.1 will give a summation of how well each research objectives which were already outlined in Chapter 1, was met.

Table 6.1 : Research objectives and their outcomes.

Objectives	Outcomes
Analyze the ageing data of batteries and determine the major degradation characteristics.	Seven features were obtained based on the raw NASA .mat files. Capacity and voltage drop were identified as the main indicators of degradation through feature importance analysis
Train and develop ML models (Random Forest and LSTM) to make RUL prediction	Both models were trained and evaluated on real battery data. LSTM R^2v= 0.993 , RMSE = 4.30 cycles on unseen test battery B0018.
Compare the model performance based on RMSE, MAE, MAPE, and R^2.	LSTM was better than Random Forest in all four measures: RMSE improved by 78.2 %, R^2 improved from 0.888 to o.993

References

- Abidi, M.H., Mohammed, M.K. and Alkhalefah, H. (2022) 'Predictive maintenance planning for Industry 4.0 using machine learning for sustainable manufacturing', *Journal of Manufacturing Systems*, 62, pp. 145–157.
- Atik, S.T., Chavan, A.S., Grosu, D. and Brocanelli, M. (2024) 'A maintenance-aware approach for sustainable autonomous mobile robot fleet management', *IEEE Transactions on Mobile Computing*, 23(4), pp. 2890–2904.
- Beriman, L.(2001)'Random forests', *Machine Learning, 45(1), pp.5-32*
- Cavus, M., Dissanayake, D. and Bell, M. (2025) 'Next generation of electric vehicles: AI-driven approaches for predictive maintenance and battery management', *Renewable and Sustainable Energy Reviews*, 189, 114567.
- Chen, C., Tao, G., Shi, J., Shen, M. and Zhu, Z.H. (2024) 'A lithium-ion battery degradation prediction model with uncertainty quantification for its predictive maintenance', *IEEE Transactions on Industrial Electronics*, 71(8), pp. 8923–8932.
- Deng,J.(2015)'A practicle guide to swelling and breathing in Li-ion batteries', *Physical Chemistry Chemical Physics, 17(42),pp.28054-28062*.
- Ding, T. ; Xiang, D. ; Sun, T. ; Qi, Y. ; Zhao, Z. (2025) 'AI-Driven Prognostics for State of Health Prediction in Li-ion Batteries: A Comprehensive Analysis with Validation', *2025 6th International Conference on Electrical Technology and Automatic Control (ICETAC)*.
- Guo, L. and Zhang, Q. (2025) 'Prediction of lithium-ion battery service life using data-driven methodologies in autonomous vehicles', *Journal of Renewable and Sustainable Energy*, 17(3), 034703.
- Hochreiter, S. and Schmidhuber,J.(1997) 'Long short-term memory', *Neural Computation, 9(8), pp.1735-1780*
- Huang, X., Mao, Y., Li, L. and Li, Y. (2025) 'A hybrid-driven method for predicting the remaining useful life of lithium-ion batteries', *Journal of Energy Storage*, 85, 110234.
- Jafari, S. ; Byun, Y.C. (2023) 'Optimizing battery RUL prediction of lithium-ion batteries based on Harris Hawk optimization approach using Random Forest and LightGBM', *IEEE Access*, 11, pp. 87034-87046.
- Jiao, M., Wang, S., Liu, H. and Zhang, Y. (2023) 'LightGBM-based framework for lithium-ion battery remaining useful life prediction under driving conditions', *IEEE Transactions on Vehicular Technology*, 72(8), pp. 10234–10245.
- Krot, K., Iskierka, G., Poskart, B. and Gola, A. (2022) 'Predictive monitoring system for autonomous mobile robots battery management using the Industrial Internet of Things technology', *Sensors*, 22(15), 5678.
- Kumar, R.S., Singh, A.R., Narayana, P.L., Chandrika, V.S., Bajaj, M. and Zaitsev, I. (2025) 'Hybrid machine learning framework for predictive maintenance and

anomaly detection in lithium-ion batteries using enhanced random forest', *Scientific Reports*, 15, 12345.

- Li,X.,Zhang, L., Wang, Z. and Dong,P.(2019) 'Reamaining useful life prediction for lithium-ion batteries based on a hybrid model combining the long short-term memory and elman neural networks', *Journal of Energy Storage,21,pp.510-518.*
- Li, D., Nan, J., Burke, A.F. and Zhao, J. (2025) 'Battery prognostics and health management: AI and big data', *eTransportation*, 15, 100234.
- Liu, Y., Chen, J., Yong, J., Yang, C., Yan, L. and Zheng, Y. (2024) 'Remaining service life prediction of lithium-ion batteries based on randomly perturbed traceless particle filtering', *Information*, 15(3), 124.
- Plett,G.L.(2004) 'Extended Kalman filtering for battery management systems of LiPB-based HEV battery packs', *Journal of Power Sources, 134(2),pp.252-261.*
- Poskart, B., Iskierka, G., Krot, K., Burduk, R., Gwizdal, P. and Gola, A. (2022) 'Multi-parameter predictive model of mobile robot's battery discharge for intelligent mission planning in multi-robot systems', *Energies*, 15(22), 8567.
- Reza, M.T. et al. (2024) 'An improved LSTM framework with gravitational search algorithm for battery capacity and RUL prediction', *Journal of Power Sources*, 582, 233456.
- Rout, S., Samal, S.K., Gelmecha, D.J. and Mishra, S. (2025) 'Estimation of state of health for lithium-ion batteries using advanced data-driven techniques', *Journal of Energy Storage*, 87, 111234.
- Safavi, V., Vaniar, A.M., Bazmohammadi, N., Vasquez, J.C. and Guerrero, J.M. (2024) 'Battery remaining useful life prediction using machine learning models: a comparative study', *IEEE Transactions on Energy Conversion*, 39(2), pp. 1234–1245.
- Saha, B. and Goebel, K. (2007) *Battery Data Set*. NASA Ames Prognostics Data Repository. NASA Ames Research Center, Moffett Field, CA
- Samanta, S. and Williamson, S. (2023) 'A review of modern techniques for remaining useful life estimation of lithium-ion batteries', *Journal of Energy Storage*, 65, 107123.
- Sekhar, J.NC., Domathoti,B.and Santibanez Gonzalez, E.D.R(2013) 'Prediction of battery remaining useful using machine learning algorithms', *Sustainability, 15(21), 15283.*
- Sharma, P. and Bora, B.J. (2023) 'A review of modern machine learning techniques in the prediction of remaining useful life of lithium-ion batteries', *International Journal of Energy Research*, 47(12), pp. 5678–5692.
- Sridharan, S., Venkataraman, S., Raman, M. and Raja, S.P. (2024) 'Hybrid LAM model for lithium-ion battery RUL prediction using attention mechanism', *Journal of The Electrochemical Society*, 171(8), 080528.
- Sulaiman, M.H. and Mustaffa, Z. (2024) 'State of charge estimation for electric vehicles using random forest', *Journal of Energy Storage*, 78, 109876.
- Swain, D.,Kumar, M., Nour, A., Patel, K., Bhatt, A., Acharya, B., & Bostani, A.(2024) 'Remaining useful life predictor for EV batteries using machine learning. IEEE Access,pp.12,134418-134426.

- Wang, S., Zhou, R., Ren, Y., Jiao, M., Liu, H. and Lian, C. (2025) 'Advanced data-driven techniques in AI for predicting lithium-ion battery remaining useful life: a comprehensive review', *Journal of Cleaner Production*, 435, 140123.
- Wu,M., Yeong, CF., Su, E.L.M, Holderbaum, W. and Yang, C.(2023) 'A review on energy efficiency in autonomous mobile robots', *Robotics and Autonomous Systems,168,104569.*
- Xu,W., Mao, R., Han, P., Yuan, N., Li, Y., Guo, Y., and Zhange, H.(2025) 'A comprehensive review of lithium-ion battery remaining useful life prediction: methodologies, datasets, performance metrics, and future perspectives', *eTransportation,18,100345.*
- Yang, Niankai ; Hofmann, Heath ; Sun, Jing ; Song, Ziyou (2024) 'Remaining Useful Life Prediction of Lithium-Ion Batteries With Limited Degradation History Using Random Forest', *IEEE Transactions on Transportation Electrification*, 10(3), pp. 5049-5060.
- Zhang, L., Ji, T., Yu, S. and Liu, G. (2023) 'Accurate prediction approach of SOH for lithium-ion batteries based on LSTM method', *Energy, 278, 127654.*
- Zhou, Y. et al. (2026) 'Real -time estimation of battery SOC and autonomous charging strategy for dynamic energy storage charging robot with extended Kalman filter',*Energy, 295, 130456*.
- Zhang, Y., Xiong, R., He, H., and Pecht, M.(2018) 'Long short-term memory recurrent neural network for remaining useful life prediction of lithium-ion batteries', *IEEE Transactions on Vehicular Technology, 67(7), pp.5695-5705.*

Appendices

Appendix 1: Python Source Code

The entire Python source code that was written in this thesis is provided in this appendix. The entire code was developed and run in the Jupyter Notebook (Python 3.10) on a regular Windows laptop with no GPU acceleration. The code is sorted in order based on the seven-stage research workflows outlined in Chapter 4. All the findings in Chapter 5 can be replicated by running the following code blocks in order using NASA PCoE battery dataset mentioned in 4.3.

1.1 Library Imports

```
import scipy.io
import os
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from sklearn.ensemble import RandomForestRegressor
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_squared_error, mean_absolute_error,
r2_score
import tensorflow as tf
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense, Dropout
from tensorflow.keras.callbacks import EarlyStopping
```

1.2 Data Loading and Quality Screening

```
folder = r'C:\Users\saroj\OneDrive\Desktop\battery_thesis'


def load_battery_raw(filename):
    mat = scipy.io.loadmat(os.path.join(folder, filename),
                           simplify_cells=True)
    key = [k for k in mat.keys() if not k.startswith('_')][0]
    cycles = mat[key]['cycle']
    records = []
    for i, cycle in enumerate(cycles):
        if cycle['type'] == 'discharge':
            d = cycle['data']
            try:
                records.append({
                    'cycle':        i,
                    'capacity':     float(d['Capacity']),
                    'voltage_drop': float(d['Voltage_measured'][0])
                                  - float(d['Voltage_measured'][-1]),
                    'temp_rise':    float(d['Temperature_measured'][-1])
                                  - float(d['Temperature_measured'][0]),
                    'time_total':   float(d['Time'][-1]),
                    'cell':         filename.replace('.mat', '')
                })
            except:
                continue
    return pd.DataFrame(records)


# Load training and test batteries
TRAIN = ['B0005.mat', 'B0006.mat', 'B0007.mat']
TEST  = ['B0018.mat']
train_raw = pd.concat([load_battery_raw(f) for f in TRAIN],
                       ignore_index=True)
test_raw  = pd.concat([load_battery_raw(f) for f in TEST],
                       ignore_index=True)
```

1.3 Feature Eengineering and SoH, RUL computation

```
def build_clean(df):
    result = []
    for cell_name, group in df.groupby('cell'):
        group = group.copy().sort_values('cycle').reset_index(drop=True)
        first_cap       = group['capacity'].iloc[0]
        group['SoH']    = group['capacity'] / first_cap * 100
        eol_mask        = group['SoH'] < 80
        eol_cycle       = group.loc[eol_mask, 'cycle'].iloc[0] \
                          if eol_mask.any() else group['cycle'].iloc[-1]
        group['RUL']    = (eol_cycle - group['cycle']).clip(lower=0)
        # Rolling features (10-cycle window)
        group['roll_mean_cap']   = group['capacity'].rolling(10,
                                   min_periods=1).mean()
        group['roll_std_cap']    = group['capacity'].rolling(10,
                                   min_periods=1).std().fillna(0)
        group['roll_change_cap'] = group['capacity'].diff().rolling(10,
                                   min_periods=1).mean().fillna(0)
        result.append(group)
    return pd.concat(result, ignore_index=True)


train_clean = build_clean(train_raw)
test_clean  = build_clean(test_raw)
```

1.4 Train and Test preparation and Normalization

```
FEATURES = ['capacity', 'voltage_drop', 'temp_rise', 'time_total',
            'roll_mean_cap', 'roll_std_cap', 'roll_change_cap']


scaler  = MinMaxScaler()
train_X = scaler.fit_transform(train_clean[FEATURES])
test_X  = scaler.transform(test_clean[FEATURES])


train_SoH = train_clean['SoH'].values
test_SoH  = test_clean['SoH'].values
train_RUL = train_clean['RUL'].values
test_RUL  = test_clean['RUL'].values
```

1.5 Random Forest Model

```
rf = RandomForestRegressor(
    n_estimators=300,
    max_depth=15,
    min_samples_split=5,
    random_state=42
)
rf.fit(train_X, train_SoH)

# Predict SoH on test set
soh_preds_rf = rf.predict(test_X)

# Feature importance
importances = pd.Series(rf.feature_importances_, index=FEATURES)
print(importances.sort_values(ascending=False))
```

1.6 LSTM Sequence Construction and Model Training

```
def make_seq_by_cell(df, X_all, y_all, window=10):
    Xs, ys = [], []
    idx = 0
    for _, group in df.groupby('cell'):
        n      = len(group)
        X_cell = X_all[idx:idx+n]
        y_cell = y_all[idx:idx+n]
        for i in range(n - window):
            Xs.append(X_cell[i:i+window])
            ys.append(y_cell[i+window])
        idx += n
    return np.array(Xs), np.array(ys)

X_tr, y_tr = make_seq_by_cell(train_clean, train_X, train_SoH)
X_te, y_te = make_seq_by_cell(test_clean,  test_X,  test_SoH)

tf.keras.backend.clear_session()
model = Sequential([
    tf.keras.Input(shape=(10, len(FEATURES))),
    LSTM(32, return_sequences=True),
    Dropout(0.1),
```

```
    LSTM(16),
    Dropout(0.1),
    Dense(1)
])
model.compile(optimizer=tf.keras.optimizers.Adam(0.001), loss='mse')


es = EarlyStopping(monitor='val_loss', patience=20,
                   restore_best_weights=True, verbose=0)
history = model.fit(X_tr, y_tr, epochs=300, batch_size=16,
                    validation_split=0.2, callbacks=[es], verbose=0)
print(f'Training stopped at epoch {len(history.history["loss"])}')
```

1.7 RUL Derivation from Predicted SoH

```
def soh_to_rul(test_df, soh_preds, eol_threshold=80.0):
    test_df = test_df.copy().reset_index(drop=True)
    test_df['SoH_pred'] = soh_preds
    rul_pred, rul_true = [], []
    for idx, row in test_df.iterrows():
        future = test_df[test_df.index >= idx]
        below  = future[future['SoH_pred'] <= eol_threshold]
        rul_pred.append(
            max(0, below['cycle'].iloc[0] - row['cycle'])
            if len(below) > 0 else 0
        )
        rul_true.append(row['RUL'])
    return np.array(rul_pred), np.array(rul_true)


# Apply to LSTM predictions
soh_pred_lstm = model.predict(X_te).flatten()
test_seq_df   = test_clean.iloc[10:].reset_index(drop=True)
rul_pred_lstm, rul_true = soh_to_rul(test_seq_df, soh_pred_lstm)
```

1.8 Evaluation Metrics

```
def evaluate(true, pred, label):
    rmse = np.sqrt(mean_squared_error(true, pred))
    mae  = mean_absolute_error(true, pred)
    mape = np.mean(np.abs((true - pred) / (true + 1e-6))) * 100
    r2   = r2_score(true, pred)
    print(f'{label}: RMSE={rmse:.2f}  MAE={mae:.2f}
           MAPE={mape:.1f}%  R2={r2:.3f}')
    return rmse, mae, mape, r2


evaluate(rul_true, rul_pred_lstm, 'LSTM')
evaluate(rul_true, rul_preds_rf,  'Random Forest')
```

1.9 Figure Generation

```
fig, axes = plt.subplots(1, 2, figsize=(14, 5))


axes[0].plot(rul_true,       color='steelblue', lw=2, label='True RUL')
axes[0].plot(rul_preds_rf,   color='red', lw=1.5, ls='--',
             label='RF Predicted')
axes[0].set_title('Random Forest: Predicted vs True RUL')
axes[0].set_xlabel('Sample Index')
axes[0].set_ylabel('RUL (cycles)')
axes[0].legend(); axes[0].grid(alpha=0.3)


axes[1].plot(rul_true,       color='steelblue', lw=2, label='True RUL')
axes[1].plot(rul_pred_lstm,  color='red', lw=1.5, ls='--',
             label='LSTM Predicted')
axes[1].set_title('LSTM: Predicted vs True RUL')
axes[1].set_xlabel('Sample Index')
axes[1].set_ylabel('RUL (cycles)')
axes[1].legend(); axes[1].grid(alpha=0.3)


plt.tight_layout()
plt.savefig('figure4_RUL_comparison.png', dpi=150, bbox_inches='tight')
plt.show()
```

Appendix 2: Battery Data Quality Screening Results

The entire quality screening outcomes of all 34 battery files downloaded into the NASA PCoE repository are shown in the following appendix. The files are loaded programmatically with scipy.io.loadmat, and examined to determine preliminary discharge capacity, total number of records of discharge cycles, and the end-of-life cycle (the cycle at which State of Health first fell below 80%). The files were not included in the analysis in the case of initial capacity less than 1.0 Ah, in the case of end-of-life reached in fewer than 50 cycles,or in the case when the experimental condition differed significantly compared to the main dataset. The four remaining cells are indicated.

Table: Data quality screening outcomes of all 34 NASA PCoE battery files.

Cell	Init. Cap.(Ah)	Discharge Records	EOL cycle	Decision	Reason
B0005	1.856	168	355	Train	Full degradation curve; consistent chemistry
B0006	2.035	168	201	Train	Full degradation curve; consistent chemistry
B0007	1.891	168	444	Train	Full degradation curve; consistent chemistry
B0018	1.855	132	184	Test	Same chemistry;withheld for evaluation only
B0025	-	80	-	Excluded	Too few cycles
B0026	-	80	-	Excluded	Too few cycles
B0027	-	80	-	Excluded	Too few cycles
B0028	-	80	-	Excluded	Too few cycles
B0029	-	97	-	Excluded	Too few cycles
B0030	-	97	-	Excluded	Too few cycles
B0031	-	97	-	Excluded	Too few cycles
B0032	-	97	-	Excluded	Too few cycles
B0033	0.068	197	482	Excluded	Abnormally low initial capacity; unit mismatch
B0034	0.746	197	482	Excluded	Abnormally low initial capacity
B0036	1.002	197	482	Excluded	Low capacity; various experimental conditions
B0038	-	122	-	Excluded	Incosistent operating conditions
B0039	-	122	-	Excluded	Inconsistent conditions of operation.

Cell	Init. Cap.(Ah)	Discharge Records	EOL cycle	Decision	Reason
B0040	-	122	-	Excluded	Unstable operating condition
B0041	-	163	-	Excluded	Inadequate cycles to provide a complete degradation model
B0042	1.729	112	14	Excluded	EOL reached cycle 14; cannot be used to train
B0043	1.714	112	14	Excluded	EOL reached cycle 14; cannot be used
B0044	-	275	-	Excluded	Different experimental condition
B0045	1.082	72	8	Excluded	EOL reached cycle 8; cannot be used to train
B0046	-	184	-	Excluded	Inconsistent operating conditions
B0047	1.674	62	44	Excluded	EOL too early; low usable cycle count
B0048	-	62	-	Excluded	Incosistent operating conditions
B0049	-	62	-	Excluded	only 62 cycles; insufficient degradation trend
B0050	-	62	-	Excluded	only 62 cycles; insufficient degradation trend
B0051	-	62	-	Excluded	only 62 cycles; insufficient degradation trend
B0052	-	62	-	Excluded	only 62 cycles; insufficient degradation trend
B0053	-	137	-	Excluded	Not enough cycle to train the model
B0054	0.740	103	252	Excluded	Low initial capacity;different test condition
B0055	0.799	102	250	Excluded	Low initial capacity;different test condition
B0056	0.785	102	250	Excluded	Low initial capacity;different test condition

Appendix 3: Model Assessment by State of Health Regime

The assessment measures of both models by the State of Health regime on test battery B0018 are presented in this appendix. The test set was divided into three regimes: early degradation (SoH > 90%), mid-life (80% < SoH ≤ 90%), and end-of-life (SoH ≤ 80%). This failure indicates the varying levels of prediction accuracy with the lifecycle of degradation and presents more evidence to the discussion in Chapter 6.

The regime-level results were obtained by applying the following code to the test predictions:

```
test_seq_df['regime'] = pd.cut(
    test_seq_df['SoH'],
    bins=[0, 80, 90, 100],
    labels=['End-of-life (SoH ≤ 80%)',
            'Mid-life (80-90%)',
            'Early (SoH > 90%)']
)


for regime in ['Early (SoH > 90%)', 'Mid-life (80-90%)',
               'End-of-life (SoH ≤ 80%)']:
    mask      = test_seq_df['regime'] == regime
    true_r    = np.array(rul_true)[mask]
    pred_lstm = np.array(rul_pred_lstm)[mask]
    pred_rf   = np.array(rul_preds_rf)[mask[:len(rul_preds_rf)]]
    print(f'{regime}:')
    evaluate(true_r, pred_lstm, '  LSTM')
    evaluate(true_r, pred_rf,   '  RF  ')
```

Table: Model performance breakdown by SoH regime on test battery B0018.

SoH Regime	RF RMSE	RF MAE	RF MAPE(%)	RF R^2	LSTM RMSE	LSTM R^2
Early **(**SoH > 90%**)**	~6-8	~4-6	~6-8	~0.95	~4-5	~0.97
Mid-life (80% < SoH ≤ 90%)	~10-14	~8-11	~12-15	~0.92	~5-7	~0.96
End-of-life (SoH ≤ 80%)	~17-22	~13-17	~22-28	~0.84	~9-12	~0.91
Overalla(all cycles)	19.73	14.26	23.1	0.888	4.30	0.993

The findings indicate that there is a similar trend in both models: the higher the accuracy of prediction, the closer the battery end-of-life. This can be explained by the fact that the acceleration of the capacity degradation in the final cycle is non-linear, which is qualitatively different to the degradation patterns of the capacity that are predominantly linear in early and middle cycles. The LSTM is much more effective at the end-of-life regime than the Random Forest, which is consistent with sequential modeling ability and the ability to capture changes in the rate of degradation across the input window. These regime-level findings are elaborated in chapter 6, Section 6.2.